

QY 61 ACAGAGATGC---AGAGAACACTCTGATGACCAAGAGACACAGAGCAGGAGAACTGG 117
117
Db 122 ACGGAGACGAGGAGAGACACCTTGGCCAGCAAGAGACCATTTGAGCANGAGAGCGG 181
118 AGTGAATTTCTTGA 132
182 AGTGAATTTCTTAA 196

RESULT 2
US-09-123-912-83
Sequence 83, Application US/09123912A
Patent No. 6312695
GENERAL INFORMATION:
APPLICANT: Reed, Steven G.
APPLICANT: Wang, Tongtong
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY OF LUNG CANCER
FILE REFERENCE: 210121.455C1
CURRENT APPLICATION NUMBER: US/09/123,912A
CURRENT FILING DATE: 1998-07-27
PRIOR APPLICATION NUMBER: 09/040,802
PRIOR FILING DATE: 1998-03-18
NUMBER OF SEQ ID NOS: 114
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 83
LENGTH: 460
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: modified_base
LOCATION: (104)
OTHER INFORMATION: Where n is a, c, g or t
NAME/KEY: modified_base
LOCATION: (118)
OTHER INFORMATION: Where n is a, c, g or t
NAME/KEY: modified_base
LOCATION: (172)
OTHER INFORMATION: Where n is a, c, g or t
NAME/KEY: modified_base
LOCATION: (401)
OTHER INFORMATION: Where n is a, c, g or t
NAME/KEY: modified_base
LOCATION: (422)
OTHER INFORMATION: Where n is a, c, g or t
NAME/KEY: modified_base
LOCATION: (423)
OTHER INFORMATION: Where n is a, c, g or t
NAME/KEY: modified_base
LOCATION: (444)
OTHER INFORMATION: Where n is a, c, g or t
NAME/KEY: modified_base
LOCATION: (449)
OTHER INFORMATION: Where n is a, c, g or t

US-09-123-912-83
Query Match 64.8%; Score 85.6; DB 4; Length 460;
Best Local Similarity 81.5%; Pred. No. 6.9e-18;
Matches 110; Conservative 0; Mismatches 22; Indels 3; Gaps 1;
QY 1 ATGGCACAACACTAGACCTGGAGAAATTCAGCTTGGATAGGCAAGCTGAAGGCC 60
117
Db 62 ATGGCAGACAAACAGACATGCGGGAATCGCCAGCTTCGATNAGGCCAAGCTGAANAA 121
61 ACAGAGATGC---AGAGAACACTCTGATGACCAAGAGACACAGAGCAGGAGAACTGG 117
122 ACGGAGACGAGGAGAGACACCTTGGCCAGCAAGAGACCATTTGAGCANGAGAGCGG 181
118 AGTGAATTTCTTGA 132
182 AGTGAATTTCTTAA 196

RESULT 3
US-09-123-912-83
Query Match 64.8%; Score 85.6; DB 4; Length 460;
Best Local Similarity 81.5%; Pred. No. 6.9e-18;
Matches 110; Conservative 0; Mismatches 22; Indels 3; Gaps 1;
QY 1 ATGGCACAACACTAGACCTGGAGAAATTCAGCTTGGATAGGCAAGCTGAAGGCC 60
117
Db 62 ATGGCAGACAAACAGACATGCGGGAATCGCCAGCTTCGATNAGGCCAAGCTGAANAA 121
61 ACAGAGATGC---AGAGAACACTCTGATGACCAAGAGACACAGAGCAGGAGAACTGG 117
122 ACGGAGACGAGGAGAGACACCTTGGCCAGCAAGAGACCATTTGAGCANGAGAGCGG 181
118 AGTGAATTTCTTGA 132
182 AGTGAATTTCTTAA 196

US-09-123-912-83
Query Match 64.8%; Score 85.6; DB 4; Length 460;
Best Local Similarity 81.5%; Pred. No. 6.9e-18;
Matches 110; Conservative 0; Mismatches 22; Indels 3; Gaps 1;
QY 1 ATGGCACAACACTAGACCTGGAGAAATTCAGCTTGGATAGGCAAGCTGAAGGCC 60
117
Db 62 ATGGCAGACAAACAGACATGCGGGAATCGCCAGCTTCGATNAGGCCAAGCTGAANAA 121
61 ACAGAGATGC---AGAGAACACTCTGATGACCAAGAGACACAGAGCAGGAGAACTGG 117
122 ACGGAGACGAGGAGAGACACCTTGGCCAGCAAGAGACCATTTGAGCANGAGAGCGG 181
118 AGTGAATTTCTTGA 132
182 AGTGAATTTCTTAA 196

US-09-123-912-83
Query Match 64.8%; Score 85.6; DB 4; Length 460;
Best Local Similarity 81.5%; Pred. No. 6.9e-18;
Matches 110; Conservative 0; Mismatches 22; Indels 3; Gaps 1;
QY 1 ATGGCACAACACTAGACCTGGAGAAATTCAGCTTGGATAGGCAAGCTGAAGGCC 60
117
Db 62 ATGGCAGACAAACAGACATGCGGGAATCGCCAGCTTCGATNAGGCCAAGCTGAANAA 121
61 ACAGAGATGC---AGAGAACACTCTGATGACCAAGAGACACAGAGCAGGAGAACTGG 117
122 ACGGAGACGAGGAGAGACACCTTGGCCAGCAAGAGACCATTTGAGCANGAGAGCGG 181
118 AGTGAATTTCTTGA 132
182 AGTGAATTTCTTAA 196

US-09-123-912-83
Query Match 64.8%; Score 85.6; DB 4; Length 460;
Best Local Similarity 81.5%; Pred. No. 6.9e-18;
Matches 110; Conservative 0; Mismatches 22; Indels 3; Gaps 1;
QY 1 ATGGCACAACACTAGACCTGGAGAAATTCAGCTTGGATAGGCAAGCTGAAGGCC 60
117
Db 62 ATGGCAGACAAACAGACATGCGGGAATCGCCAGCTTCGATNAGGCCAAGCTGAANAA 121
61 ACAGAGATGC---AGAGAACACTCTGATGACCAAGAGACACAGAGCAGGAGAACTGG 117
122 ACGGAGACGAGGAGAGACACCTTGGCCAGCAAGAGACCATTTGAGCANGAGAGCGG 181
118 AGTGAATTTCTTGA 132
182 AGTGAATTTCTTAA 196

US-09-123-912-83
Query Match 64.8%; Score 85.6; DB 4; Length 460;
Best Local Similarity 81.5%; Pred. No. 6.9e-18;
Matches 110; Conservative 0; Mismatches 22; Indels 3; Gaps 1;
QY 1 ATGGCACAACACTAGACCTGGAGAAATTCAGCTTGGATAGGCAAGCTGAAGGCC 60
117
Db 62 ATGGCAGACAAACAGACATGCGGGAATCGCCAGCTTCGATNAGGCCAAGCTGAANAA 121
61 ACAGAGATGC---AGAGAACACTCTGATGACCAAGAGACACAGAGCAGGAGAACTGG 117
122 ACGGAGACGAGGAGAGACACCTTGGCCAGCAAGAGACCATTTGAGCANGAGAGCGG 181
118 AGTGAATTTCTTGA 132
182 AGTGAATTTCTTAA 196

US-09-643-597-83
Sequence 83, Application US/09643597
Patent No. 6426072
GENERAL INFORMATION:
APPLICANT: Wang, Tongtong
APPLICANT: Fan, Liqun
APPLICANT: Kalos, Michael D.
APPLICANT: Bangur, Chaitanya S.
APPLICANT: Hosken, Nancy
APPLICANT: Fauger, Gary R.
APPLICANT: Li, Samuel X.
APPLICANT: Wang, Aijun
APPLICANT: Skeiky, Yasir A.W.
APPLICANT: Henderson, Robert A.
APPLICANT: McNeill, Patricia D.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
FILE REFERENCE: 210121.455C11
CURRENT APPLICATION NUMBER: US/09/643,597
CURRENT FILING DATE: 2000-08-21
NUMBER OF SEQ ID NOS: 369
SOFTWARE: FastSEQ for Windows Version 3.0
SEQ ID NO 83
LENGTH: 460
TYPE: DNA
ORGANISM: Homo sapien
FEATURE:
NAME/KEY: misc.feature
LOCATION: (1)--(460)
OTHER INFORMATION: n = A,T,C or G

US-09-643-597-83
Query Match 64.8%; Score 85.6; DB 4; Length 460;
Best Local Similarity 81.5%; Pred. No. 6.9e-18;
Matches 110; Conservative 0; Mismatches 22; Indels 3; Gaps 1;
QY 1 ATGGCACAACACTAGACCTGGAGAAATTCAGCTTGGATAGGCAAGCTGAAGGCC 60
117
Db 62 ATGGCAGACAAACAGACATGCGGGAATCGCCAGCTTCGATNAGGCCAAGCTGAANAA 121
61 ACAGAGATGC---AGAGAACACTCTGATGACCAAGAGACACAGAGCAGGAGAACTGG 117
122 ACGGAGACGAGGAGAGACACCTTGGCCAGCAAGAGACCATTTGAGCANGAGAGCGG 181
118 AGTGAATTTCTTGA 132
182 AGTGAATTTCTTAA 196

US-09-643-597-83
Query Match 64.8%; Score 85.6; DB 4; Length 460;
Best Local Similarity 81.5%; Pred. No. 6.9e-18;
Matches 110; Conservative 0; Mismatches 22; Indels 3; Gaps 1;
QY 1 ATGGCACAACACTAGACCTGGAGAAATTCAGCTTGGATAGGCAAGCTGAAGGCC 60
117
Db 62 ATGGCAGACAAACAGACATGCGGGAATCGCCAGCTTCGATNAGGCCAAGCTGAANAA 121
61 ACAGAGATGC---AGAGAACACTCTGATGACCAAGAGACACAGAGCAGGAGAACTGG 117
122 ACGGAGACGAGGAGAGACACCTTGGCCAGCAAGAGACCATTTGAGCANGAGAGCGG 181
118 AGTGAATTTCTTGA 132
182 AGTGAATTTCTTAA 196

US-09-643-597-83
Query Match 64.8%; Score 85.6; DB 4; Length 460;
Best Local Similarity 81.5%; Pred. No. 6.9e-18;
Matches 110; Conservative 0; Mismatches 22; Indels 3; Gaps 1;
QY 1 ATGGCACAACACTAGACCTGGAGAAATTCAGCTTGGATAGGCAAGCTGAAGGCC 60
117
Db 62 ATGGCAGACAAACAGACATGCGGGAATCGCCAGCTTCGATNAGGCCAAGCTGAANAA 121
61 ACAGAGATGC---AGAGAACACTCTGATGACCAAGAGACACAGAGCAGGAGAACTGG 117
122 ACGGAGACGAGGAGAGACACCTTGGCCAGCAAGAGACCATTTGAGCANGAGAGCGG 181
118 AGTGAATTTCTTGA 132
182 AGTGAATTTCTTAA 196

US-09-643-597-83
Query Match 64.8%; Score 85.6; DB 4; Length 460;
Best Local Similarity 81.5%; Pred. No. 6.9e-18;
Matches 110; Conservative 0; Mismatches 22; Indels 3; Gaps 1;
QY 1 ATGGCACAACACTAGACCTGGAGAAATTCAGCTTGGATAGGCAAGCTGAAGGCC 60
117
Db 62 ATGGCAGACAAACAGACATGCGGGAATCGCCAGCTTCGATNAGGCCAAGCTGAANAA 121
61 ACAGAGATGC---AGAGAACACTCTGATGACCAAGAGACACAGAGCAGGAGAACTGG 117
122 ACGGAGACGAGGAGAGACACCTTGGCCAGCAAGAGACCATTTGAGCANGAGAGCGG 181
118 AGTGAATTTCTTGA 132
182 AGTGAATTTCTTAA 196

US-09-643-597-83
Query Match 64.8%; Score 85.6; DB 4; Length 460;
Best Local Similarity 81.5%; Pred. No. 6.9e-18;
Matches 110; Conservative 0; Mismatches 22; Indels 3; Gaps 1;
QY 1 ATGGCACAACACTAGACCTGGAGAAATTCAGCTTGGATAGGCAAGCTGAAGGCC 60
117
Db 62 ATGGCAGACAAACAGACATGCGGGAATCGCCAGCTTCGATNAGGCCAAGCTGAANAA 121
61 ACAGAGATGC---AGAGAACACTCTGATGACCAAGAGACACAGAGCAGGAGAACTGG 117
122 ACGGAGACGAGGAGAGACACCTTGGCCAGCAAGAGACCATTTGAGCANGAGAGCGG 181
118 AGTGAATTTCTTGA 132
182 AGTGAATTTCTTAA 196

US-09-643-597-83
Query Match 64.8%; Score 85.6; DB 4; Length 460;
Best Local Similarity 81.5%; Pred. No. 6.9e-18;
Matches 110; Conservative 0; Mismatches 22; Indels 3; Gaps 1;
QY 1 ATGGCACAACACTAGACCTGGAGAAATTCAGCTTGGATAGGCAAGCTGAAGGCC 60
117
Db 62 ATGGCAGACAAACAGACATGCGGGAATCGCCAGCTTCGATNAGGCCAAGCTGAANAA 121
61 ACAGAGATGC---AGAGAACACTCTGATGACCAAGAGACACAGAGCAGGAGAACTGG 117
122 ACGGAGACGAGGAGAGACACCTTGGCCAGCAAGAGACCATTTGAGCANGAGAGCGG 181
118 AGTGAATTTCTTGA 132
182 AGTGAATTTCTTAA 196

US-09-643-597-83
Query Match 64.8%; Score 85.6; DB 4; Length 460;
Best Local Similarity 81.5%; Pred. No. 6.9e-18;
Matches 110; Conservative 0; Mismatches 22; Indels 3; Gaps 1;
QY 1 ATGGCACAACACTAGACCTGGAGAAATTCAGCTTGGATAGGCAAGCTGAAGGCC 60
117
Db 62 ATGGCAGACAAACAGACATGCGGGAATCGCCAGCTTCGATNAGGCCAAGCTGAANAA 121
61 ACAGAGATGC---AGAGAACACTCTGATGACCAAGAGACACAGAGCAGGAGAACTGG 117
122 ACGGAGACGAGGAGAGACACCTTGGCCAGCAAGAGACCATTTGAGCANGAGAGCGG 181
118 AGTGAATTTCTTGA 132
182 AGTGAATTTCTTAA 196

US-09-643-597-83
Query Match 64.8%; Score 85.6; DB 4; Length 460;
Best Local Similarity 81.5%; Pred. No. 6.9e-18;
Matches 110; Conservative 0; Mismatches 22; Indels 3; Gaps 1;
QY 1 ATGGCACAACACTAGACCTGGAGAAATTCAGCTTGGATAGGCAAGCTGAAGGCC 60
117
Db 62 ATGGCAGACAAACAGACATGCGGGAATCGCCAGCTTCGATNAGGCCAAGCTGAANAA 121
61 ACAGAGATGC---AGAGAACACTCTGATGACCAAGAGACACAGAGCAGGAGAACTGG 117
122 ACGGAGACGAGGAGAGACACCTTGGCCAGCAAGAGACCATTTGAGCANGAGAGCGG 181
118 AGTGAATTTCTTGA 132
182 AGTGAATTTCTTAA 196

Best Local Similarity 66.7%; Pred. No. 9.9e-05;
Matches 78; Conservative 0; Mismatches 36; Indels 3; Gaps 1;
QY 1 ATGCGACACAACTAGACCTGGAGAAATTCGCCAGCTTGGTAAGGCCCAAGCC 60
Db 78 ATGCTGACAAACCCGATGGCTGAGATCGAGAAATTCGATAAGTGAAGAAAG 137
QY 61 ACAGAGATGCGAGAGAACTCTGTATGA---CCAAAGAGACACAGAGAGAGAG 114
Db 138 ACAGAGACGCAAGAGAAATCCACTGCCTTCCAAAGAAAGATTTGAACAGAGAG 194

RESULT 5

US-09-058-489-26
; Sequence 26, Application US/09058489
; Patent No. 6103886
; GENERAL INFORMATION:
; APPLICANT: Whitehead Institute for Biomedical Research
; APPLICANT: Lahn, Bruce
; APPLICANT: Page, David
; TITLE OF INVENTION: Genes in the No. 6103886-Recombining Region of
; FILE REFERENCE: WHI97-08PA
; CURRENT APPLICATION NUMBER: US/09/058.489
; CURRENT FILING DATE: 1998-04-10
; EARLIER APPLICATION NUMBER: 60/041.877
; EARLIER FILING DATE: 1997-04-11
; NUMBER OF SEQ ID NOS: 91
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 26
; LENGTH: 1161
; TYPE: DNA
; ORGANISM: Human
US-09-058-489-26

Query Match 29.2%; Score 38.6; DB 3; Length 1161;
Best Local Similarity 64.1%; Pred. No. 0.004;
Matches 75; Conservative 0; Mismatches 39; Indels 3; Gaps 1;

QY 1 ATGCGACACAACTAGACCTGGAGAAATTCGCCAGCTTGGTAAGGCCCAAGCC 60
Db 768 ATGCTGACAAACCTGGTATGGCTGAGATCGAGAAATTCGATAAGTGAAGAG 827
QY 61 ACAGAGATGCGAGAGAACTCTGTATGA---CCAAAGAGACACAGAGAGAGAG 114
Db 828 ACAGAAACGCAAGAGAAATCCATTTGCTTCCAAAGAAACTATCGAAAGAGAG 884

RESULT 6

US-08-751-782-11
; Sequence 11, Application US/08751782
; Patent No. 5821352
; GENERAL INFORMATION:
; APPLICANT: Heintz, Nathaniel
; APPLICANT: Gubbay, Johnathan
; APPLICANT: Skinner, Michael
; TITLE OF INVENTION: A CDNA Library Prepared during
; TITLE OF INVENTION: Regression of Rat Prostate and Methods of Use Thereof
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: David A. Jackson, Esq.
; STREET: 411 Hackensack Ave, Continental Plaza, 4th
; STREET: Floor
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/751,782
; FILING DATE: 18-NOV-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Jackson Esq., David A.
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 600-1-190
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-487-5800
; TELEFAX: 201-343-1684
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 182 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; IMMEDIATE SOURCE:
; CLONE: 10.27
US-08-751-782-11

Query Match 27.6%; Score 36.4; DB 1; Length 182;
Best Local Similarity 64.5%; Pred. No. 0.01;
Matches 71; Conservative 0; Mismatches 36; Indels 3; Gaps 1;
QY 8 ACAAACTAGACCTGGAAGAAATTCGAGCTTGGTAAGGCCCAAGCCACAGAGA 67
Db 1 ACAAACCCGATATGGCTGAGATCGAGAAATTCGATAAGTGAAGAGAGAGAGAA 60
QY 68 TGC---AGAAGAACACTCTGTATGACCAAGAGACACACAGAGAGAGAGAG 114
Db 61 CACAAGAGAAAAATCTCTCTCTCAAAAGAAACAAATTCGAACAAGAGAGAG 110

RESULT 7

US-08-925-171-11
; Sequence 11, Application US/08925171
; Patent No. 5928871
; GENERAL INFORMATION:
; APPLICANT: Heintz, Nathaniel
; APPLICANT: Gubbay, Johnathan
; APPLICANT: Skinner, Michael
; TITLE OF INVENTION: A CDNA Library Prepared during
; TITLE OF INVENTION: Regression of Rat Prostate and Methods of Use Thereof
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: David A. Jackson, Esq.
; STREET: 411 Hackensack Ave, Continental Plaza, 4th
; STREET: Floor
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/925,171
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Jackson Esq., David A.
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 600-1-190 CIP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-487-5800
; TELEFAX: 201-343-1684
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 182 base pairs

```

TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: CDNA
HYPOTHETICAL: NO
IMMEDIATE SOURCE:
CLONE: 10.27
US-08-925-171-11

Query Match      27.6%; Score 36.4; DB 2; Length 182;
Best Local Similarity 64.5%; Pred. No. 0.01;
Matches 71; Conservative 0; Mismatches 36; Indels 3; Gaps 1;

QY 8 ACAACATAGACTCGAAGAATAATCCCACTTGGTAAGGCCAACGTCGAAGCCACACAGAGA 67
    ||||| || || || || || || || || || || || || || || || || || || || ||
Db 1 ACAAAACCCGATATGCCTCAGATCGAGAAAATTCGATAAGTGTGAAGTTGAAGAAGACAGAAA 60
    ||||| || || || || || || || || || || || || || || || || || || || ||

QY 68 TGC---AGAAGAACACATCTGATGACCAGAACAGACACACAGACGAGGAGAA 114
    | |||| || |||| || |||| || |||| || |||| || |||| || |||| || ||||
Db 61 CACAAGAGAAAACTCTGCTCTCAAAGAGAAACAATTGAACAGAGAAG 110
    | |||| || |||| || |||| || |||| || |||| || |||| || |||| || ||||

RESULT 8
US-09-484-970B-2
; Sequence 2, Application US/09484970B
; Patent No. 6426186
; GENERAL INFORMATION:
; APPLICANT: Jones, Karen A.
; APPLICANT: Volkmar, Wayne
; TITLE OF INVENTION: BONE REMODELING GENES
; FILE REFERENCE: PB-0014 US
; CURRENT APPLICATION NUMBER: US/09/484.970B
; CURRENT FILING DATE: 2000-01-18
; NUMBER OF SEQ ID NOS: 172
; SOFTWARE: PERL Program
; SEQ ID NO 2
; LENGTH: 4117
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. 6426186 344594 (1794154CB1)
US-09-484-970B-2

Query Match      23.8%; Score 31.4; DB 4; Length 4117;
Best Local Similarity 61.7%; Pred. No. 1.1;
Matches 50; Conservative 0; Mismatches 31; Indels 0; Gaps 0;

QY 44 AGGCCAAGCTGAAGGCCACAGAGATGCAGAGAACACATCTGATCACCAAGAACACACAG 103
    || |||| || |||| || |||| || |||| || |||| || |||| || |||| || ||||
Db 2558 ATGAAAAGCTTGAGAAACACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2617
    || |||| || |||| || |||| || |||| || |||| || |||| || |||| || ||||

QY 104 AGCAGAGAGACGTGGAGTGA 124
    || || || || || || || || || || || || || || || || || || || || ||
Db 2618 ACAAGTTACTTAAGAGTGA 2638
    || || || || || || || || || || || || || || || || || || || || ||

RESULT 9
US-08-145-705A-1/c
; Sequence 1, Application US/08145705A
; Patent No. 5489513
; GENERAL INFORMATION:
; APPLICANT: Springer, Wolfgang; Piempel, Manfred;
; APPLICANT: Lberding, Antonius
; TITLE OF INVENTION: SPECIFIC GENE PROBES AND
; TITLE OF INVENTION: PROCESSES FOR THE DIAGNOSTIC
; TITLE OF INVENTION: INVESTIGATION OF CANDIDA
; TITLE OF INVENTION: ALBICANS
; NUMBER OF SEQUENCES: 44
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SPRUNG HORN KRAMER & WOODS

```



```
APPLICATION NUMBER: US/08/232.463
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/935.313
FILING DATE:
APPLICATION NUMBER: EP 91 114 300.6
FILING DATE: 26-AUG-1991
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 30472/114 IMM
TELEPHONE: (703)836-9300
TELEFAX: (703)683-4109
TELEX: 899149
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 7218 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
CLONE: pTZgpt-Fls
US-08-232-463-14

Query Match 23.3%; Score 30.8; DB 1; Length 7218;
Best Local Similarity 12.7%; Pred. No. 2;
Matches 14; Conservative 62; Mismatches 34; Indels 0; Gaps 0;

Qy 15 AGACTGGAGAAATGCGACCTGGATPAGCCAGCTGAAGCCACAGACAGATCGAGAA 74
||| ||||| ||| : : : : : : : : : : : : : : : : : :
Db 1455 AGATAGAGAAATTTGTACRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1396

Qy 75 GAACACTGTGATGACCAAGAGACACAGACAGGAGAGAGTGGAGTGA 124
: : : : : : : : : : : : : : : : : : : : : : : :
Db 1395 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1346

RESULT 11
US-08-688-376-1
Sequence 1, Application US/08688376
Patent No. 6018039
GENERAL INFORMATION:
APPLICANT: Satow, Hiroyasu
TITLE OF INVENTION: NOVEL PROCESS FOR PRODUCING SUBSTANCES
TITLE OF INVENTION: IN MAMMARY GLAND OF TRANSGENIC ANIMAL BY USING MC26 GENE
TITLE OF INVENTION: EXPRESSION-REGULATORY REGION
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: DILWORTH & BARRESE
STREET: 4350 LaJolla Village Drive, Suite 300
CITY: San Diego
STATE: CA
COUNTRY: USA
ZIP: 92122
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/688.376
FILING DATE: 30-JUL-1996
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: Pepper, Frederick W.
REGISTRATION NUMBER: 31,286
REFERENCE/DOCKET NUMBER: 567-3
TELEPHONE: 619-546-4410
TELEFAX: 619-453-2839
INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:
LENGTH: 5394 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CAAT_signal
LOCATION: 2234..2243
FEATURE:
NAME/KEY: TATA_signal
LOCATION: 2275..2281
FEATURE:
NAME/KEY: polyA_signal
LOCATION: 4607..4612
US-08-688-376-1

Query Match 23.2%; Score 30.6; DB 3; Length 5394;
Best Local Similarity 58.1%; Pred. No. 2.1;
Matches 54; Conservative 0; Mismatches 39; Indels 0; Gaps 0;

Qy 39 GGATAGGCCCAAGCTGAAGCCACAGACAGATGCGAGAGAACACTCTGTATGACCAAGAGAC 98
||| ||||| ||| : : : : : : : : : : : : : : : : : :
Db 1445 GAAGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1504

Qy 99 CACAGAGCAGGAGAGAGTGGAGTGAAATTTCTCTG 131
||| ||||| ||| : : : : : : : : : : : : : : : : : :
Db 1505 GAAGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1537

RESULT 12
US-08-145-705A-3/c
Sequence 3, Application US/08145705A
Patent No. 5489513
GENERAL INFORMATION:
APPLICANT: Springer, Wolfgang; Plempel, Manfred;
APPLICANT: L bberding, Antonius
TITLE OF INVENTION: SPECIFIC GENE PROBES AND
TITLE OF INVENTION: PROCESSES FOR THE DIAGNOSTIC
TITLE OF INVENTION: INVESTIGATION OF CANDIDA
TITLE OF INVENTION: ALBICANS
NUMBER OF SEQUENCES: 44
CORRESPONDENCE ADDRESS:
ADDRESSEE: SPRUNG HORN KRAMER & WOODS
STREET: 660 White Plains Road
CITY: Tarrytown
STATE: New York
COUNTRY: U.S.A.
ZIP: 10591-5144
MEDIUM TYPE: Diskette, 3.5 inch, 1.4 MB storage
COMPUTER: NEC PowerMate 1 Plus
OPERATING SYSTEM: DOS
SOFTWARE: WordPerfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/145,705A
FILING DATE: October 28, 1993
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: German P 42 36 708.5
FILING DATE: October 30, 1992
ATTORNEY/AGENT INFORMATION:
NAME: Kurt G. Briscoe
REGISTRATION NUMBER: 33,141
REFERENCE/DOCKET NUMBER: Bayer 8885-KGB
TELEPHONE: (914) 332-1700
TELEFAX: (914) 332-1844
TELEX:
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 100 base pairs
TYPE: nucleic acid
```

STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Candida albicans
US-08-145-705A-3

Query Match 23.0%; Score 30.4; DB 1; Length 100;
Best Local Similarity 61.2%; Pred. No. 0.61; Mismatches 0; Gaps 0;
Matches 49; Conservative 0; Indels 31; Indels 0; Gaps 0;
QY 44 AGCCAGAGCTGAAGGCCACAGATGCGAAGAACACTCTGATGACCAAGAGACACAG 103
DB 97 AGCCCAAGAAGAGCGAGAGGCCCAAGAAGAGGAGGAGGCCCAAGAAGAGGAGCAG 38
QY 104 AGCAGGAGAGTGGAGTGA 123
DB 37 AGGAGGCCAAGAAGAGGCA 18

RESULT 13
US-08-742-185-101
Sequence 101, Application US/08742185
Patent No. 6020476
GENERAL INFORMATION:
APPLICANT: Page, David C.
APPLICANT: Reijo, Renee
APPLICANT: Saxena, Richa
APPLICANT: Hawkins, Trevor
APPLICANT: Reeve, Mary Pat
TITLE OF INVENTION: DAZ: A GENE FAMILY ASSOCIATED WITH AZOOSPERMIA
NUMBER OF SEQUENCES: 102
CORRESPONDENCE ADDRESS:
ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
STREET: Two Militia Drive
CITY: Lexington
STATE: Massachusetts
COUNTRY: US
ZIP: 02173

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08742.185
FILING DATE: 30-OCT-1996
CLASSIFICATION: 435

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/690,734
FILING DATE: 31-JUL-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/310,429
FILING DATE: 22-SEP-1994

ATTORNEY/AGENT INFORMATION:
NAME: Granahan, Patricia
REGISTRATION NUMBER: 32,227
REFERENCE/DOCKET NUMBER: WHI94-07A2
TELEPHONE: (617) 861-6240
TELEFAX: (617) 861-9540

INFORMATION FOR SEQ ID NO: 101:
SEQUENCE CHARACTERISTICS:
LENGTH: 43795 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-742-185-101

Query Match 22.7%; Score 30; DB 3; Length 43795;

Best Local Similarity 59.3%; Pred. No. 6.6;
Matches 51; Conservative 0; Mismatches 35; Indels 0; Gaps 0;
QY 39 GGATAAGGCCAAGCTGAAGGCCACAGATGCGAAGAACACTCTGATGACCAAGAGAC 98
DB 38962 GAAGAGGAG 39021
QY 99 CACAGAGCAG 124
DB 39022 GAGGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGA 39047

RESULT 14
US-08-664-856A-1
Sequence 1, Application US/08664856A
Patent No. 5663071
GENERAL INFORMATION:
APPLICANT: BRUCE R. ZETTER AND LERE BAO
TITLE OF INVENTION: HUMAN THYMOSIN 15 GENE,
TITLE OF INVENTION: PROTEIN AND USES THEREOF
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN
STREET: 130 WATER STREET
CITY: BOSTON
STATE: MA
COUNTRY: USA
ZIP: 02019

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/664,856A
FILING DATE: 17 JUN 1996
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:

ATTORNEY/AGENT INFORMATION:
NAME: DAVID, RESNICK S
REGISTRATION NUMBER: 34,235
REFERENCE/DOCKET NUMBER: 46507
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-523-3400
TELEFAX: 617-523-6440
TELEX: 200291 STRE
INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:
LENGTH: 412 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE:
ORIGINAL SOURCE:
FEATURE:
NAME/KEY: Coding Sequence
LOCATION: 98...232
OTHER INFORMATION: Exon 1
US-08-664-856A-1

Query Match 22.4%; Score 29.6; DB 1; Length 412;
Best Local Similarity 59.5%; Pred. No. 1.8;
Matches 69; Conservative 0; Mismatches 44; Indels 3; Gaps 1;

QY 1 ATGGCACACAAACTAGACCTGGAGAAATTCAGCTTGGATGAAGCCAGCTGAAGGCC 60
DB 98 ATGAGTGATAAACACCAGACTTATCAGAAGTTGAACATTTGACAAATCAAGATTGAAGAAG 157

THIS PAGE BLANK (13PT0)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 2, 2003, 13:24:07 ; Search time 214 Seconds
(without alignments)
1389.083 Million cell updates/sec

Title: US-09-915-178-1

Perfect score: 132

Sequence: 1 atggcacacaaactagacct.....agtggagtgaaatttcctga 132

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 2185239 segs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_101002.*

- 1: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
- 2: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
- 3: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
- 4: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
- 5: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
- 6: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
- 7: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
- 8: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
- 9: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
- 10: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
- 11: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
- 12: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
- 13: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
- 14: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
- 15: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
- 16: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
- 17: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
- 18: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
- 19: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
- 20: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
- 21: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
- 22: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
- 23: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
- 24: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	132	100.0	147	24	Human NOV2a encodi
2	132	100.0	147	24	Human NOV2b encodi
3	132	100.0	147	24	Human NOV2c encodi
4	88.6	67.1	243	24	Human colon cancer
5	88.6	67.1	353	24	Human colon cancer
6	88.6	67.1	395	24	Human colon cancer
7	88.6	67.1	428	24	Human colon cancer
8	88.6	67.1	429	24	Human colon cancer
9	88.6	67.1	434	24	Human colon cancer

10	88.6	67.1	438	24	ABK54988	Human colon cancer
11	88.6	67.1	438	24	ABK55113	Human colon cancer
12	88.6	67.1	439	24	ABK55440	Human colon cancer
13	88.6	67.1	443	24	ABK55189	Human colon cancer
14	88.6	67.1	445	24	ABK53257	Human colon cancer
15	88.6	67.1	446	24	ABK55421	Human colon cancer
16	88.6	67.1	451	24	ABK55083	Human colon cancer
17	88.6	67.1	452	24	ABK54878	Human colon cancer
18	88.6	67.1	453	24	ABN96522	Gene #3020 used to
19	88.6	67.1	453	24	ABK54804	Human colon cancer
20	88.6	67.1	455	24	ABK54803	Human colon cancer
21	88.6	67.1	456	24	ABK54749	Human colon cancer
22	88.6	67.1	465	24	ABK54586	Human colon cancer
23	88.6	67.1	488	21	AAC10524	Human secreted pro
24	88.6	67.1	491	22	AAD03640	Human thymosin b-1
25	88.6	67.1	517	24	ABK55404	Human colon cancer
26	88.6	67.1	534	23	ABV33875	Human prostate exp
27	88.6	67.1	534	23	ABV42758	Human prostate exp
28	88.6	67.1	693	24	ABK55328	Human colon cancer
29	88.6	67.1	836	23	ABV12746	Human prostate exp
30	88.6	67.1	1201	23	ABV21155	Human prostate exp
31	88.6	67.1	1201	23	ABV22339	Human prostate exp
32	88.6	67.1	1201	23	ABV28163	Human prostate exp
33	88	66.7	430	22	AAS00689	CDNA encoding thym
34	88	66.7	430	24	AAD38157	Thymosin-beta-10-1
35	85.6	64.8	460	20	AAC224583	Human lung tumor a
36	85.6	64.8	460	21	AAC65822	Human lung cancer-
37	85.6	64.8	460	24	ABL49041	Human lung tumour
38	82.2	62.3	446	24	ABK55170	Human colon cancer
39	80.8	61.2	267	21	AAC00908	Human secreted pro
40	77.8	58.9	449	24	ABK54652	Human colon cancer
41	77.8	58.9	451	24	ABK54720	Human colon cancer
42	73.2	55.5	292	23	ABV03577	Human prostate exp
43	70.2	53.2	763	20	AZ27510	Human ovarian tumo
44	63.4	48.0	453	21	AAC10525	Human secreted pro
45	63.2	47.9	724	24	AAS62054	Porcine muscular s

ALIGNMENTS

RESULT 1
ABA92649
ID ABA92649 standard; CDNA; 147 BP.
XX AC ABA92649;
XX DT 25-MAR-2002 (first entry)
XX DE Human NOV2a encoding polynucleotide SEQ ID NO:3.
XX DE Human; NOVX; cytostatic; antidiabetic; anorectic; antibacterial;
KW fungicide; virucide; protozoacide; analgesic; antiparkinsonian;
KW antilasthmatic; hypotensive; osteopathic; antiinflammatory; antitumor;
KW neuroprotective; cardiatic; antiallergic; antidepressant; nootropic;
KW anticonvulsant; neuroleptic; gene therapy; vaccine; immune disease;
KW developmental disease; taste and scent detectability disorder; infection;
KW Burkitt's lymphoma; signal transduction pathway disorder; pain; cancer;
KW retinal disease; feeding disorder; asthma; Parkinson's disease; ulcer;
KW noninsulin-dependent diabetes mellitus; acute heart failure; hypotension;
KW hypertension; urinary retention; osteoporosis; Crohn's disease; allergy;
KW multiple sclerosis; angina pectoris; myocardial infarction; delirium;
KW benign prostatic hypertrophy; manic depression; dementia; dyskinesia;
KW severe mental retardation; Huntington's disease; gene; chromosome 9;
KW Gilles de la Tourette syndrome; ss.

Homo sapiens.
Key Location/Qualifiers
CDS 6..137
/*tag= a
/product= "NOV2a"

	Matches	132;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0;
QY	1	ATGGCACACAAACTAGACCTTGGGAAGAAATTCGCCAGTTGGATAGGCCAACAGCTGAAGGCC	60							
Db	6	ATGGCACACAAACTAGACCTTGGGAAGAAATTCGCCAGTTGGATAGGCCAACAGCTGAAGGCC	65							
QY	61	ACAGAGATGCAGGAAGAACACTCTGTGATGACCAAGAGACCACAGAGCAGGAGAGACTGGAGT	120							
Db	66	ACAGAGATGCAGGAAGAACACTCTGTGATGACCAAGAGACCACAGAGCAGGAGAGACTGGAGT	125							
QY	121	GAATTTTCTCGA	132							
Db	126	GAATTTTCTCGA	137							

RESULT 4	
ABK55047	
ID	ABK55047 standard; cDNA; 243 BP.
XX	
AC	ABK55047;
XX	
DT	18-JUN-2002 (first entry)
XX	
DE	Human colon cancer-associated cDNA, SEQ ID No 517.
XX	
KW	Human; colon cancer; immunogenic; vaccine; tumour; gene; ss.
XX	
OS	Homo sapiens.
XX	
PN	WO200212280-A2.
XX	
PD	14-FEB-2002.
XX	
PF	30-JUL-2001; 2001WO-US23826.
XX	
PR	03-AUG-2000; 2000US-223265P.
PR	02-OCT-2000; 2000US-237406P.
PR	20-MAR-2001; 2001US-277495P.
PR	03-JUL-2001; 2001US-302702P.
XX	
PA	(CORI-) CORIYA CORP.
XX	
PI	Pyle RA, Xu J, Secrist H;
XX	
DR	WPI; 2002-257462/30.
XX	
PT	Novel polynucleotide encoding colon tumour polypeptides, useful as
PT	vaccines for treating colon cancers -
XX	
PS	Claim 1; Page 297; 425pp; English.
XX	
CC	The invention relates to isolated polynucleotides (I) encoding colon
CC	tumour polypeptides (II). (I) is useful for stimulating an immune
CC	response in a patient and treating colon cancer in a patient.
CC	Oligonucleotides derived from (I) are useful for determining the presence
CC	of cancer in a patient. (I) and (II) are useful in pharmaceutical
CC	compositions, e.g. vaccines, and other compositions for the diagnosis
CC	and treatment of colon cancer. A composition comprising a first component
CC	selected from physiologically acceptable carriers and immunostimulants,
CC	and an antigen-presenting cell expressing (II) is useful for inhibiting
CC	development of cancer in a patient. (I) is useful in the design and
CC	preparation of ribozyme molecules for inhibiting expression of tumour
CC	polypeptides and (I). ABK54531-ABK55464 represent human colon cancer cDNA
CC	sequences of the invention.
XX	
SO	sequence 243 BP; 73 A; 64 C; 68 G; 37 T; 1 other;

```

Query Match      67.1%; Score 88.6; DB 24; Length 243;
Best Local Similarity 83.7%; Pred. No. 2e-16;
Matches 113; Conservative 0; Mismatches 19; Indels 3; Gaps 1;

QY      1  ATGGCACAAACTAGACCTCGAAGAAATTCGCAGTTGGATAGGCCAAGCTCAAGGCC 60
      |||||

```

Db	63	ATGGCAGACAAACAGACATGGGGAAATCGCCAGCTTCGATAGGCCCAAGCTGAAGAAA	122
Qy	61	ACAGAGATGC-- --AGAGAACACTCTCATCAACCAAGACAGACAGCAGGAGAAGTGG	117
Db	123	ACGGAGACGCGAGAGAAGAACACCCCTGCCGACCACCAAGAGACCAATTGACACAGGAGACGG	182
Qy	118	AGTGAATTTTCCTGA	132
Db	183	AGTGAATTTTCCTAA	197

RESULT 5
 BK54754
 BK54754

RESULT 5	ABK34754	ABK54754 standard; cDNA; 353 BP.
ID	ABK54754	standard; cDNA; 353 BP.
XX	ABK54754;	
XX	18-JUN-2002	(first entry)
DT	XX	Human colon cancer-associated cDNA, SEQ
DE	DE	Human; colon cancer; immunogenic; vacci
KW	KW	Human; colon cancer; immunogenic; vacci
XX	OS	Homo sapiens.
XX	XX	WO200212280-A2.
XX	XX	14-FEB-2002.
XX	XX	30-JUL-2001; 2001WO-US23826.
PF	XX	03-AUG-2000; 2000US-223265P.
XX	XX	02-OCT-2000; 2000US-237406P.
PR	PR	20-MAR-2001; 2001US-277495P.
PR	PR	03-JUL-2001; 2001US-302702P.
XX	XX	(CORI-) CORIXA CORP.
PA	PA	Pyle RA, Xu J, Secretist H;
XX	PI	WPI; 2002-257462/30.
XX	XX	Novel polynucleotide encoding colon tum
PT	PT	vaccines for treating colon cancers -
PT	PT	Claim 1; Page 214; 425pp; English.
XX	PS	The invention relates to isolated polyn
XX	CC	tumour polypeptides (II). (I) is useful
CC	CC	response in a patient and treating col
CC	CC	Oligonucleotides derived from (I) are
CC	CC	of cancer in a patient. (I) and (II) at
CC	CC	compositions, e.g. vaccines, and other
CC	CC	and treatment of colon cancer. A compos
CC	CC	selected from physiologically acceptab
CC	CC	and an antigen-presenting cell express
CC	CC	development of cancer in a patient. (I
CC	CC	preparation of ribozyme molecules for
CC	CC	polypeptides and (I). ABK54531-ABK5546
CC	CC	sequences of the invention.
XX	XX	Sequence 353 BP; 97 A; 103 C; 98 G;-55
XX	XX	Best Match 67.1%; Score 88.
XX	XX	Query local Similarity 83.7%; Pred. No.

QY 1 ATGGCACAAACTAGACTTGAAGAAATTGCCAGCTGGGATAGGCCAAGCTGAAGCC 60

Db 37 ATGGCAGACAAACCAGACATGGGGNAATCGCCAGCTTCGATAAGSCCAAGCTGAAGAA 96

QY 61 ACAGAGATGC---AGAAGAACACTCTGTATGACCAAGAGAGACCACAGCAGGAGAAAGTGG 117


```
Db 97 ACGGAGCAGCAGGAGAGAACACCCCTGCGGACCAAGAGACCATTTGAGCAGGAGAGCGG 156
QY 118 AGTGAATTTCTCTGA 132
    |||||
Db 157 AGTGAATTTCTCTAA 171

RESULT 6
ABK54737
ID ABK54737 standard; cDNA; 395 BP.
XX
AC ABK54737;
XX
DT 18-JUN-2002 (first entry)
XX
DE Human colon cancer-associated cDNA, SEQ ID No 207.
XX
DE Human colon cancer; immunogenic; vaccine; tumour; gene; ss.
XX
KW Homo sapiens.
XX
OS WO200212280-A2.
XX
PN 14-FEB-2002.
XX
PD 30-JUL-2001; 2001WO-US23826.
XX
PF 03-AUG-2000; 2000US-223265P.
XX
PR 02-OCT-2000; 2000US-237406P.
XX
PR 20-MAR-2001; 2001US-277495P.
XX
PR 03-JUL-2001; 2001US-302702P.
XX
PA (CORI-) CORIXA CORP.
XX
PI Pyle RA, Xu J, Secrist H;
XX
XX WPI; 2002-257462/30.
XX
PT Novel polynucleotide encoding colon tumour polypeptides, useful as
XX vaccines for treating colon cancers -
XX
PS Claim 1; Page 209; 425pp; English.
XX
CC The invention relates to isolated polynucleotides (I) encoding colon
XX tumour polypeptides (II). (I) is useful for stimulating an immune
XX response in a patient and treating colon cancer in a patient.
XX CC Oligonucleotides derived from (I) are useful for determining the presence
XX of cancer in a patient. (I) and (II) are useful in pharmaceutical
XX compositions, e.g. vaccines, and other compositions for the diagnosis
XX and treatment of colon cancer. A composition comprising a first component
XX selected from physiologically acceptable carriers and immunostimulants,
XX and an antigen-presenting cell expressing (II) is useful for inhibiting
XX development of cancer in a patient. (I) is useful in the design and
XX preparation of ribozyme molecules for inhibiting expression of tumour
XX polypeptides and (I). ABK54531-ABK55464 represent human colon cancer
XX sequences of the invention.
XX
SQ Sequence 395 BP; 105 A; 113 C; 115 G; 62 T; 0 other;

Query Match 67.1%; Score 88.6; DB 24; Length 395;
Best Local Similarity 83.7%; Pred. No. 2.3e-16;
Matches 113; Conservative 0; Mismatches 19; Indels 3; Gaps 1;

QY 1 ATGGCAGACAACTAGACCTTGGAGAAATTCGACGTTGGATAGGCCAAGCTGAAGGCC 60
    |||||
Db 72 ATGGCAGACAAACACGACATCGGGGAAATTCGACGTTGGATAGGCCAAGCTGAAGAAA 131
    |||||

QY 61 ACAGAGATGC---AGAAGAACACTCTGATGACCAAGAGACCAACAGAGAGGAGAGTGG 117
    |||||
Db 132 ACGGAGCAGCAGGAGAGAGACACCCCTGCCGACCAAGAGACCATTTGAGCAGGAGAGCGG 191
    |||||

QY 118 AGTGAATTTCTCTGA 132
    |||||
```

```
Db 192 AGTGAATTTCTCTAA 206

RESULT 7
ABK55033
ID ABK55033 standard; cDNA; 428 BP.
XX
AC ABK55033;
XX
DT 18-JUN-2002 (first entry)
XX
DE Human colon cancer-associated cDNA, SEQ ID No 503.
XX
DE Human colon cancer; immunogenic; vaccine; tumour; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200212280-A2.
XX
PD 14-FEB-2002.
XX
PF 30-JUL-2001; 2001WO-US23826.
XX
PR 03-AUG-2000; 2000US-223265P.
XX
PR 02-OCT-2000; 2000US-237406P.
XX
PR 20-MAR-2001; 2001US-277495P.
XX
PR 03-JUL-2001; 2001US-302702P.
XX
PA (CORI-) CORIXA CORP.
XX
PI Pyle RA, Xu J, Secrist H;
XX
XX WPI; 2002-257462/30.
XX
PT Novel polynucleotide encoding colon tumour polypeptides, useful as
XX vaccines for treating colon cancers -
XX
PS Claim 1; Page 294; 425pp; English.
XX
CC The invention relates to isolated polynucleotides (I) encoding colon
XX tumour polypeptides (II). (I) is useful for stimulating an immune
XX response in a patient and treating colon cancer in a patient.
XX CC Oligonucleotides derived from (I) are useful for determining the presence
XX of cancer in a patient. (I) and (II) are useful in pharmaceutical
XX compositions, e.g. vaccines, and other compositions for the diagnosis
XX and treatment of colon cancer. A composition comprising a first component
XX selected from physiologically acceptable carriers and immunostimulants,
XX and an antigen-presenting cell expressing (II) is useful for inhibiting
XX development of cancer in a patient. (I) is useful in the design and
XX preparation of ribozyme molecules for inhibiting expression of tumour
XX polypeptides and (I). ABK54531-ABK55464 represent human colon cancer
XX sequences of the invention.
XX
SQ Sequence 428 BP; 124 A; 114 C; 115 G; 75 T; 0 other;

Query Match 67.1%; Score 88.6; DB 24; Length 428;
Best Local Similarity 83.7%; Pred. No. 2.4e-16;
Matches 113; Conservative 0; Mismatches 19; Indels 3; Gaps 1;

QY 1 ATGGCAGACAACTAGACCTTGGAGAAATTCGACGTTGGATAGGCCAAGCTGAAGGCC 60
    |||||
Db 50 ATGGCAGACAAACACGACATCGGGGAAATTCGACGTTGGATAGGCCAAGCTGAAGAAA 109
    |||||

QY 61 ACAGAGATGC---AGAAGAACACTCTGATGACCAAGAGACCAACAGAGAGGAGAGTGG 117
    |||||
Db 110 ACGGAGCAGCAGGAGAGAGACACCCCTGCCGACCAAGAGACCATTTGAGCAGGAGAGCGG 169
    |||||

QY 118 AGTGAATTTCTCTGA 132
    |||||
Db 170 AGTGAATTTCTCTAA 184

RESULT 8
```

ABK55163
ID ABK55163 standard; cDNA; 429 BP.

XX AC ABK55163;
XX DT 18-JUN-2002 (first entry)

XX DE Human colon cancer-associated cDNA, SEQ ID No 633.

XX KW Human; colon cancer; immunogenic; vaccine; tumour; gene; ss.
XX OS Homo sapiens.
XX PN WO200212280-A2.
XX PD 14-FEB-2002.

XX PF 30-JUL-2001; 2001WO-US23826.
XX PR 03-AUG-2000; 2000US-223265P.
XX PR 02-OCT-2000; 2000US-237406P.
XX PR 20-MAR-2001; 2001US-277495P.
XX PR 03-JUL-2001; 2001US-302702P.

XX PA (CORI-) CORIXA CORP.
XX PI Pyle RA, Xu J, Secrlist H;
XX DR WPI; 2002-257462/30.

XX PT Novel polynucleotide encoding colon tumour polypeptides, useful as
vaccines for treating colon cancers -
XX PS Claim 1; Page 329; 425pp; English.
XX CC The invention relates to isolated polynucleotides (I) encoding colon
tumour polypeptides (II). (I) is useful for stimulating an immune
response in a patient and treating colon cancer in a patient.
XX CC Oligonucleotides derived from (I) are useful for determining the presence
of cancer in a patient. (I) and (II) are useful in pharmaceutical
compositions, e.g. vaccines, and other compositions for the diagnosis
and treatment of colon cancer. A composition comprising a first component
selected from physiologically acceptable carriers and immunostimulants,
and an antigen-presenting cell expressing (II) is useful for inhibiting
development of cancer in a patient. (I) is useful in the design and
preparation of ribozyme molecules for inhibiting expression of tumour
polypeptides and (I). ABK55163-ABK55164 represent human colon cancer cDNA
sequences of the invention.

XX SQ Sequence 429 BP; 124 A; 115 C; 115 G; 75 T; 0 other;

Query Match 67.1%; Score 88.6; DB 24; Length 429;
Best Local Similarity 83.7%; Pred. No. 2.4e-16;
Matches 113; Conservative 0; Mismatches 19; Indels 3; Gaps 1;
QY 1 ATGGCACAACACTAGACCTGGAGAAATTCGACCTTGGATAGGCCCAAGCTGAAGGCC 60
DB 51 ATGGCACAACACTAGACCTGGAGAAATTCGACCTTGGATAGGCCCAAGCTGAAGAAA 110
QY 61 ACAGAGATGC---AGAAGAACACTCTGTATGACCAAGAGACACAGAGAGAGAGTGG 117
DB 111 ACGGAGACGAGGAGAGAGACACCTCTGCCCAACCAAGAGACCATTTGAGCAGGAGAGCGG 170
QY 118 AGTGAATTTCTCTGA 132
DB 171 AGTGAATTTCTCTAA 185

RESULT 9
ABK55157
ID ABK55157 standard; cDNA; 434 BP.

XX AC ABK55157;
XX DT 18-JUN-2002 (first entry)

XX DE Human colon cancer-associated cDNA, SEQ ID No 458.

XX DT 18-JUN-2002 (first entry)

XX DE Human colon cancer-associated cDNA, SEQ ID No 627.

XX KW Human; colon cancer; immunogenic; vaccine; tumour; gene; ss.

XX OS Homo sapiens.
XX PN WO200212280-A2.
XX PD 14-FEB-2002.

XX PF 30-JUL-2001; 2001WO-US23826.
XX PR 03-AUG-2000; 2000US-223265P.
XX PR 02-OCT-2000; 2000US-237406P.
XX PR 20-MAR-2001; 2001US-277495P.
XX PR 03-JUL-2001; 2001US-302702P.

XX PA (CORI-) CORIXA CORP.
XX PI Pyle RA, Xu J, Secrlist H;
XX DR WPI; 2002-257462/30.

XX PT Novel polynucleotide encoding colon tumour polypeptides, useful as
vaccines for treating colon cancers -
XX PS Claim 1; Page 328; 425pp; English.
XX CC The invention relates to isolated polynucleotides (I) encoding colon
tumour polypeptides (II). (I) is useful for stimulating an immune
response in a patient and treating colon cancer in a patient.
XX CC Oligonucleotides derived from (I) are useful for determining the presence
of cancer in a patient. (I) and (II) are useful in pharmaceutical
compositions, e.g. vaccines, and other compositions for the diagnosis
and treatment of colon cancer. A composition comprising a first component
selected from physiologically acceptable carriers and immunostimulants,
and an antigen-presenting cell expressing (II) is useful for inhibiting
development of cancer in a patient. (I) is useful in the design and
preparation of ribozyme molecules for inhibiting expression of tumour
polypeptides and (I). ABK55163-ABK55164 represent human colon cancer cDNA
sequences of the invention.

XX SQ Sequence 434 BP; 124 A; 116 C; 117 G; 77 T; 0 other;

Query Match 67.1%; Score 88.6; DB 24; Length 434;
Best Local Similarity 83.7%; Pred. No. 2.4e-16;
Matches 113; Conservative 0; Mismatches 19; Indels 3; Gaps 1;
QY 1 ATGGCACAACACTAGACCTGGAGAAATTCGACCTTGGATAGGCCCAAGCTGAAGGCC 60
DB 54 ATGGCACAACACTAGACCTGGAGAAATTCGACCTTGGATAGGCCCAAGCTGAAGAAA 113
QY 61 ACAGAGATGC---AGAAGAACACTCTGTATGACCAAGAGACACAGAGAGAGTGG 117
DB 114 ACGGAGACGAGGAGAGAGACACCTCTGCCCAAGAGACCATTTGAGCAGGAGAGCGG 173
QY 118 AGTGAATTTCTCTGA 132
DB 174 AGTGAATTTCTCTAA 188

RESULT 10
ABK54988
ID ABK54988 standard; cDNA; 438 BP.

XX AC ABK54988;
XX DT 18-JUN-2002 (first entry)

XX DE Human colon cancer-associated cDNA, SEQ ID No 458.

XX Human; colon cancer; immunogenic; vaccine; tumour; gene; ss.
 XX Homo sapiens.
 OS

PN WO200212280-A2.

XX 14-FEB-2002.

XX 30-JUL-2001; 2001WO-US23826.

XX 03-AUG-2000; 2000US-223265P.

PR 02-OCT-2000; 2000US-237406P.

PR 20-MAR-2001; 2001US-277495P.

PR 03-JUL-2001; 2001US-302702P.

XX (CORI-) CORIXA CORP.

PA Pyle RA, Xu J, Secretist H;

XX WPI; 2002-257462/30.

XX Novel polynucleotide encoding colon tumour polypeptides, useful as
 PT vaccines for treating colon cancers -
 PT
 XX

PS Claim 1; Page 280-281; 425pp; English.

XX The invention relates to isolated polynucleotides (I) encoding colon
 CC tumour polypeptides (II). (I) is useful for stimulating an immune
 CC response in a patient and treating colon cancer in a patient.
 CC Oligonucleotides derived from (I) are useful for determining the presence
 CC of cancer in a patient. (I) and (II) are useful in pharmaceutical
 CC compositions, e.g. vaccines, and other compositions for the diagnosis
 CC and treatment of colon cancer. A composition comprising a first component
 CC selected from physiologically acceptable carriers and immunostimulants,
 CC and an antigen-presenting cell expressing (II) is useful for inhibiting
 CC development of cancer in a patient. (I) is useful in the design and
 CC preparation of ribozyme molecules for inhibiting expression of tumour
 CC polypeptides and (I). ABK54531-ABK55464 represent human colon cancer cDNA
 CC sequences of the invention.

XX Sequence 438 BP; 126 A; 118 C; 117 G; 77 T; 0 other;

Query Match 67.1%; Score 88.6; DB 24; Length 438;
 Best Local Similarity 83.7%; Pred. No. 2.4e-16;
 Matches 113; Conservative 0; Mismatches 19; Indels 3; Gaps 1;

QY 1 ATGGCACAACACTAGACCTGGAGAAATTCGACGCTTGATAGGCCCAAGCTGAAGGCC 60
 Db 60 ATGGCAGACAAACACAGACATGGGGAAATCGCCAGCTTCGATAGGCCCAAGCTGAAGAAA 119
 QY 61 ACAGAGATGC---AGAGAACACTCTGATGACCAAGAGACCACAGAGCAGGAGAGTGG 117
 Db 120 ACGGAGACCGAGGAGAGAACACCCCTGCCGACCAAGAGACCATTGAGCAGGAGAGCGG 179
 QY 118 AGTGAATTTCTCTGA 132
 Db 180 AGTGAATTTCTCTAA 194

RESULT 11

ABK55113

ID ABK55113 standard; cDNA; 438 BP.

XX ABK55113;

XX 18-JUN-2002 (first entry)

DT Human colon cancer-associated cDNA, SEQ ID No 583.

XX Human; colon cancer; immunogenic; vaccine; tumour; gene; ss.

XX Homo sapiens.

XX WO200212280-A2.

XX 14-FEB-2002.

XX 30-JUL-2001; 2001WO-US23826.

XX 03-AUG-2000; 2000US-223265P.

PR 02-OCT-2000; 2000US-237406P.

PR 20-MAR-2001; 2001US-277495P.

PR 03-JUL-2001; 2001US-302702P.

XX (CORI-) CORIXA CORP.

XX Pyle RA, Xu J, Secretist H;

XX WPI; 2002-257462/30.

XX Novel polynucleotide encoding colon tumour polypeptides, useful as
 PT vaccines for treating colon cancers -
 PT
 XX

PS Claim 1; Page 315-316; 425pp; English.

XX The invention relates to isolated polynucleotides (I) encoding colon
 CC tumour polypeptides (II). (I) is useful for stimulating an immune
 CC response in a patient and treating colon cancer in a patient.
 CC Oligonucleotides derived from (I) are useful for determining the presence
 CC of cancer in a patient. (I) and (II) are useful in pharmaceutical
 CC compositions, e.g. vaccines, and other compositions for the diagnosis
 CC and treatment of colon cancer. A composition comprising a first component
 CC selected from physiologically acceptable carriers and immunostimulants,
 CC and an antigen-presenting cell expressing (II) is useful for inhibiting
 CC development of cancer in a patient. (I) is useful in the design and
 CC preparation of ribozyme molecules for inhibiting expression of tumour
 CC polypeptides and (I). ABK54531-ABK55464 represent human colon cancer cDNA
 CC sequences of the invention.

XX Sequence 438 BP; 126 A; 118 C; 117 G; 77 T; 0 other;

Query Match 67.1%; Score 88.6; DB 24; Length 438;
 Best Local Similarity 83.7%; Pred. No. 2.4e-16;
 Matches 113; Conservative 0; Mismatches 19; Indels 3; Gaps 1;

QY 1 ATGGCACAACACTAGACCTGGAGAAATTCGACGCTTGATAGGCCCAAGCTGAAGGCC 60
 Db 60 ATGGCAGACAAACACAGACATGGGGAAATCGCCAGCTTCGATAGGCCCAAGCTGAAGAAA 119
 QY 61 ACAGAGATGC---AGAGAACACTCTGATGACCAAGAGACCACAGAGCAGGAGAGTGG 117
 Db 120 ACGGAGACCGAGGAGAGAACACCCCTGCCGACCAAGAGACCATTGAGCAGGAGAGCGG 179
 QY 118 AGTGAATTTCTCTGA 132
 Db 180 AGTGAATTTCTCTAA 194

RESULT 12

ABK55440

ID ABK55440 standard; cDNA; 439 BP.

XX ABK55440;

XX 18-JUN-2002 (first entry)

DT Human colon cancer-associated cDNA, SEQ ID No 910.

XX Human; colon cancer; immunogenic; vaccine; tumour; gene; ss.

XX Homo sapiens.

XX WO200212280-A2.

XX 14-FEB-2002.

PA (CORI-) CORIXA CORP.

XX Pyle RA, Xu J, Secríst H;

XX DR WPI; 2002-257462/30.

XX Novel polynucleotide encoding colon tumour polypeptides, useful as
PT vaccines for treating colon cancers

XX Claim 1; Page 355; 425pp; English.

XX The invention relates to isolated polynucleotides (I) encoding colon
CC tumour polypeptides (II). (I) is useful for stimulating an immune
CC response in a patient and treating colon cancer in a patient.
CC Oligonucleotides derived from (I) are useful for determining the presence
CC of cancer in a patient. (I) and (II) are useful in pharmaceutical
CC compositions, e.g. vaccines, and other compositions for the diagnosis
CC and treatment of colon cancer. A composition comprising a first component
CC selected from physiologically acceptable carriers and immunostimulants,
CC and an antigen-presenting cell expressing (II) is useful for inhibiting
CC development of cancer in a patient. (I) is useful in the design and
CC preparation of ribozyme molecules for inhibiting expression of tumour
CC polypeptides and (I). ABK54531-ABK55464 represent human colon cancer cDNA
CC sequences of the invention.

XX Sequence 445 BP; 130 A; 119 C; 119 G; 77 T; 0 other;

Query Match 67.1%; Score 88.6; DB 24; Length 445;

Best Local Similarity 83.7%; Pred. No. 2.4e-16;

Matches 113; Conservative 0; Mismatches 19; Indels 3; Gaps 1;

QY 1 ATGGCACAACTAGACCTGGGAAGAAATTCGAGCTTGGATAGGCCCAAGCTGAAGGCC 60

Db ATGGCACAACTAGACCTGGGAAGAAATTCGAGCTTGGATAGGCCCAAGCTGAAGAAA 119

QY 61 ACAGAGATGC---AGAAGAACACTCTGATGACCAAGAGACACAGAGAGAGAGTGG 117

Db 120 ACGGAGACGAGGAGAGAACACCTCGCGACCAAGAGACCATTCGAGCAGGAGAGCGG 179

QY 118 AGTGAATTTCTCTGA 132

Db 180 AGTGAATTTCTCTAA 194

RESULT 15

ABK55421

ID ABK55421 standard; cDNA; 446 BP.

XX AC ABK55421;

XX DT 18-JUN-2002 (first entry)

XX DE Human colon cancer-associated cDNA, SEQ ID NO 891.

XX KW Human; colon cancer; immunogenic; vaccine; tumour; gene; ss.

XX OS Homo sapiens.

XX PN WO200212280-A2.

XX PD 14-FEB-2002.

XX PF 30-JUL-2001; 2001WO-US23826.

XX PR 03-AUG-2000; 2000US-223265P.

XX PR 02-OCT-2000; 2000US-237406P.

XX PR 20-MAR-2001; 2001US-277495P.

XX PR 03-JUL-2001; 2001US-302702P.

XX PA (CORI-) CORIXA CORP.

XX PI Pyle RA, Xu J, Secríst H;

XX

DR WPI; 2002-257462/30.

XX Novel polynucleotide encoding colon tumour polypeptides, useful as
PT vaccines for treating colon cancers

XX Claim 1; Page 408; 425pp; English.

XX The invention relates to isolated polynucleotides (I) encoding colon
CC tumour polypeptides (II). (I) is useful for stimulating an immune
CC response in a patient and treating colon cancer in a patient.
CC Oligonucleotides derived from (I) are useful for determining the presence
CC of cancer in a patient. (I) and (II) are useful in pharmaceutical
CC compositions, e.g. vaccines, and other compositions for the diagnosis
CC and treatment of colon cancer. A composition comprising a first component
CC selected from physiologically acceptable carriers and immunostimulants,
CC and an antigen-presenting cell expressing (II) is useful for inhibiting
CC development of cancer in a patient. (I) is useful in the design and
CC preparation of ribozyme molecules for inhibiting expression of tumour
CC polypeptides and (I). ABK54531-ABK55464 represent human colon cancer cDNA
CC sequences of the invention.

XX Sequence 446 BP; 141 A; 113 C; 111 G; 79 T; 2 other;

Query Match 67.1%; Score 88.6; DB 24; Length 446;

Best Local Similarity 83.7%; Pred. No. 2.4e-16;

Matches 113; Conservative 0; Mismatches 19; Indels 3; Gaps 1;

QY 1 ATGGCACAACTAGACCTGGGAAGAAATTCGAGCTTGGATAGGCCCAAGCTGAAGGCC 60

Db 44 ATGGCACAACTAGACCTGGGAAGAAATTCGAGCTTGGATAGGCCCAAGCTGAAGAAA 103

QY 61 ACAGAGATGC---AGAAGAACACTCTGATGACCAAGAGACACAGAGAGAGTGG 117

Db 104 ACGGAGACGAGGAGAGAACACCTCGCGACCAAGAGACCATTCGAGCAGGAGAGCGG 163

QY 118 AGTGAATTTCTCTGA 132

Db 164 AGTGAATTTCTCTAA 178

Search completed: June 2, 2003, 14:18:23

Job time : 215 secs

THIS PAGE BLANK (38PT0)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 2, 2003, 14:18:32 ; Search time 106 Seconds
(without alignments)
1679.096 Million cell updates/sec

Title: US-09-915-178-1

Perfect score: 132

Sequence: 1 atggcacacaaactagacct.....agtggagtgaatttcctga 132

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 845702 seqs, 674182571 residues

Total number of hits satisfying chosen parameters: 1691404

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications_NA.*

- 1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
- 2: /cgn2_6/ptodata/2/pubpna/ECT_NEW_PUB.seq.*
- 3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
- 4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*
- 5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
- 6: /cgn2_6/ptodata/2/pubpna/ECTUS_PUBCOMB.seq.*
- 7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
- 8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq.*
- 9: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
- 10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq.*
- 11: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq.*
- 12: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq.*
- 13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*
- 14: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Length	ID	Description
1	88.6	67.1	243	10	US-09-919-580-517
2	88.6	67.1	353	10	US-09-919-580-224
3	88.6	67.1	395	10	US-09-919-580-207
4	88.6	67.1	426	10	US-09-919-580-71
5	88.6	67.1	428	10	US-09-919-580-503
6	88.6	67.1	429	10	US-09-919-580-633
7	88.6	67.1	434	10	US-09-919-580-627
8	88.6	67.1	438	10	US-09-919-580-458
9	88.6	67.1	438	10	US-09-919-580-583
10	88.6	67.1	439	10	US-09-919-580-910
11	88.6	67.1	443	10	US-09-919-580-659
12	88.6	67.1	445	10	US-09-919-580-727
13	88.6	67.1	446	10	US-09-919-580-891
14	88.6	67.1	449	9	US-09-918-995-15768
15	88.6	67.1	451	10	US-09-919-580-553
16	88.6	67.1	452	10	US-09-919-580-348
17	88.6	67.1	453	10	US-09-919-580-274
18	88.6	67.1	453	10	US-09-880-107-3019
19	88.6	67.1	455	10	US-09-919-580-273

20	88.6	67.1	456	10	US-09-919-580-219
21	88.6	67.1	465	10	US-09-919-580-56
22	88.6	67.1	468	9	US-09-918-995-21161
23	88.6	67.1	469	9	US-09-918-995-24062
24	88.6	67.1	473	9	US-09-918-995-15273
25	88.6	67.1	483	9	US-09-918-995-19906
26	88.6	67.1	517	10	US-09-919-580-874
27	88.6	67.1	593	10	US-09-919-580-798
28	85.6	64.8	460	10	US-09-735-705-83
29	85.6	64.8	460	10	US-09-850-7168-83
30	85.6	64.8	460	10	US-09-897-778-83
31	83.8	63.5	461	9	US-09-918-995-22450
32	82.2	62.3	446	10	US-09-919-580-640
33	80.6	61.1	465	9	US-09-918-995-23067
34	80.2	60.8	494	9	US-09-918-995-24178
35	77.8	58.9	449	10	US-09-919-580-122
36	77.8	58.9	451	10	US-09-919-580-190
37	77.8	58.9	474	9	US-09-918-995-22477
38	57.4	43.5	559	10	US-09-919-580-159
39	54.4	41.2	446	10	US-09-880-107-3158
40	50.2	38.0	404	10	US-09-960-352-10885
41	50.2	38.0	405	10	US-09-960-352-6906
42	50.2	38.0	408	10	US-09-960-352-12808
43	50.2	38.0	461	10	US-09-960-352-2605
44	46.6	35.3	468	9	US-09-918-995-13407
45	45.4	34.4	377	10	US-09-960-352-6901

ALIGNMENTS

RESULT 1
US-09-919-580-517
; Sequence 517, Application US/09919580
; Patent No. US20020110832A1
; GENERAL INFORMATION:
; APPLICANT: Pyle, Ruth
; APPLICANT: Xu, Jiangchun
; APPLICANT: Secrist, Heather
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF COLON CANCER
; FILE REFERENCE: 210121.552
; CURRENT APPLICATION NUMBER: US/09/919,580
; CURRENT FILING DATE: 2001-07-30
; NUMBER OF SEQ ID NOS: 934
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 517
; LENGTH: 243
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 243
; OTHER INFORMATION: n = A,T,C or G
US-09-919-580-517

Query Match	67.1%	Score 88.6;	DB 10;	Length 243;
Best Local Similarity	83.7%	Pred. No. 4.6e-19;		
Matches 113;	Conservative 0;	Mismatches 19;	Indels 3;	Gaps 1;
QY	1	ATGCGACACAACTAGACCTGGAGAAATTCGCCAGTTGGATAGGCCAAGCTGAAGGCC	60	
Db	63	ATGCGACACAACTAGACCTGGAGAAATTCGCCAGTTGGATAGGCCAAGCTGAAGGCC	122	
QY	61	ACAGAGATGC---AGAGACACATCTGATGACCAAGAGACACACAGCAGCAGAGTGG	117	
Db	123	ACGAGACGCGAGGAGAGACACCCCTGCCGACCAAGAGACCATTTGACGAGGAGCGG	182	
QY	118	AGTGAATTTCTCTGA 132		
Db	183	AGTGAATTTCTCTAA 197		

```
RESULT 2
US-09-919-580-224
; Sequence 224, Application US/09919580
; Patent No. US20020110832A1
; GENERAL INFORMATION:
; APPLICANT: Pyle, Ruth
; APPLICANT: Xu, Jiangchun
; APPLICANT: Secretist, Heather
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF COLON CANCER
; FILE REFERENCE: 210121.552
; CURRENT APPLICATION NUMBER: US/09/919,580
; CURRENT FILING DATE: 2001-07-30
; NUMBER OF SEQ ID NOS: 934
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 224
; LENGTH: 353
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-919-580-224

Query Match      67.1%; Score 88.6; DB 10; Length 353;
Best Local Similarity 83.7%; Pred. No. 5.1e-19;
Matches 113; Conservative 0; Mismatches 19; Indels 3; Gaps 1;

QY 1 ATGGCACACAACCTAGACCTGGAGAAATTCGCCAGCTTGGATAGGCGCAAGCTGAAGGCC 60
Db 37 ATGGCAGACAACCCAGACATGGGGGAAATCGCCAGCTTCGATAGGCGCAAGCTGAAGAAA 96
QY 61 ACAGAGATGC---AGAAGAACAACCTCTGATGACCAAGAGACACACAGCAGGAGAGTGG 117
Db 97 ACGGAGCGCAGGAGAGAACACCTCGCCGACCAAGAGACCATTTGACGAGGAGAGCGG 156
QY 118 AGTGAATTTCTCTGA 132
Db 157 AGTGAATTTCTCTAA 171

RESULT 3
US-09-919-580-207
; Sequence 207, Application US/09919580
; Patent No. US20020110832A1
; GENERAL INFORMATION:
; APPLICANT: Pyle, Ruth
; APPLICANT: Xu, Jiangchun
; APPLICANT: Secretist, Heather
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF COLON CANCER
; FILE REFERENCE: 210121.552
; CURRENT APPLICATION NUMBER: US/09/919,580
; CURRENT FILING DATE: 2001-07-30
; NUMBER OF SEQ ID NOS: 934
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 207
; LENGTH: 395
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-919-580-207

Query Match      67.1%; Score 88.6; DB 10; Length 395;
Best Local Similarity 83.7%; Pred. No. 5.2e-19;
Matches 113; Conservative 0; Mismatches 19; Indels 3; Gaps 1;

QY 1 ATGGCACACAACCTAGACCTGGAGAAATTCGCCAGCTTGGATAGGCGCAAGCTGAAGGCC 60
Db 72 ATGGCAGACAACCCAGACATGGGGGAAATTCGCCAGCTTCGATAGGCGCAAGCTGAAGAAA 131
QY 61 ACAGAGATGC---AGAAGAACAACCTCTGATGACCAAGAGACACACAGCAGGAGAGTGG 117
Db 132 ACGGAGCGCAGGAGAGAACACCTCGCCGACCAAGAGACCATTTGACGAGGAGAGCGG 191
QY 118 AGTGAATTTCTCTGA 132
Db 157 AGTGAATTTCTCTAA 171

RESULT 4
US-09-998-598-71
; Sequence 71, Application US/09998598
; Patent No. US20020150922A1
; GENERAL INFORMATION:
; APPLICANT: Stolk, John A.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Chenault, Ruth A.
; APPLICANT: Meagher, Madelein Joy
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF COLON CANCER
; FILE REFERENCE: 210121.561
; CURRENT APPLICATION NUMBER: US/09/998,598
; CURRENT FILING DATE: 2001-11-16
; NUMBER OF SEQ ID NOS: 2606
; SOFTWARE: Corixa Invention Disclosure Database
; SEQ ID NO 71
; LENGTH: 426
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-998-598-71

Query Match      67.1%; Score 88.6; DB 10; Length 426;
Best Local Similarity 83.7%; Pred. No. 5.3e-19;
Matches 113; Conservative 0; Mismatches 19; Indels 3; Gaps 1;

QY 1 ATGGCACACAACCTAGACCTGGAGAAATTCGCCAGCTTGGATAGGCGCAAGCTGAAGGCC 60
Db 48 ATGGCAGACAACCCAGACATGGGGGAAATTCGCCAGCTTCGATAGGCGCAAGCTGAAGAAA 107
QY 61 ACAGAGATGC---AGAAGAACAACCTCTGATGACCAAGAGACACACAGCAGGAGAGTGG 117
Db 108 ACGGAGCGCAGGAGAGAACACCTCGCCGACCAAGAGACCATTTGACGAGGAGAGCGG 167
QY 118 AGTGAATTTCTCTGA 132
Db 168 AGTGAATTTCTCTAA 182

RESULT 5
US-09-919-580-503
; Sequence 503, Application US/09919580
; Patent No. US20020110832A1
; GENERAL INFORMATION:
; APPLICANT: Pyle, Ruth
; APPLICANT: Xu, Jiangchun
; APPLICANT: Secretist, Heather
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF COLON CANCER
; FILE REFERENCE: 210121.552
; CURRENT APPLICATION NUMBER: US/09/919,580
; CURRENT FILING DATE: 2001-07-30
; NUMBER OF SEQ ID NOS: 934
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 503
; LENGTH: 428
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-919-580-503

Query Match      67.1%; Score 88.6; DB 10; Length 428;
Best Local Similarity 83.7%; Pred. No. 5.3e-19;
Matches 113; Conservative 0; Mismatches 19; Indels 3; Gaps 1;

QY 1 ATGGCACACAACCTAGACCTGGAGAAATTCGCCAGCTTGGATAGGCGCAAGCTGAAGGCC 60
Db 50 ATGGCAGACAACCCAGACATGGGGGAAATTCGCCAGCTTCGATAGGCGCAAGCTGAAGAAA 109
QY 61 ACAGAGATGC---AGAAGAACAACCTCTGATGACCAAGAGACACACAGCAGGAGAGTGG 117
Db 157 AGTGAATTTCTCTAA 171
```


Db 110 ACGGAGACGACGAGGAAGAACACCCCTGCCGACCAAGAGACCACTTGCAGCAGGAGAAGCGG 169
QY 118 AGTGAATTTCTCTGA 132
Db 170 AGTGAATTTCTCTAA 184

RESULT 6

US-09-919-580-633
; Sequence 633, Application US/09919580
; Patent No. US20020110832A1

GENERAL INFORMATION:

; APPLICANT: Pyle, Ruth
; APPLICANT: Xu, Jiangchun
; APPLICANT: Secrist, Heather
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF COLON CANCER
; FILE REFERENCE: 210121.552
; CURRENT APPLICATION NUMBER: US/09/919,580
; CURRENT FILING DATE: 2001-07-30
; NUMBER OF SEQ ID NOS: 934
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 633
; LENGTH: 429
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-919-580-633

Query Match 67.1%; Score 88.6; DB 10; Length 429;
Best Local Similarity 83.7%; Pred. No. 5.3e-19;
Matches 113; Conservative 0; Mismatches 19; Indels 3; Gaps 1;

QY 1 ATGGCAGACAACTAGACCTGGGAAGAAATTCAGCTTGGATAAGGCCAAGCTGAAGGCC 60
Db 51 ATGGCAGACAACTAGACCTGGGAAGAAATTCAGCTTGGATAAGGCCAAGCTGAAGAAA 110
QY 61 ACAGAGATGC---AGAAGAACACTCTGATGACCAAGAGACCAACAGCAGGAGAGAGTGG 117
Db 111 ACGGAGACGACGAGGAAGAACACCCCTGCCGACCAAGAGACCACTTGCAGCAGGAGAAGCGG 170
QY 118 AGTGAATTTCTCTGA 132
Db 171 AGTGAATTTCTCTAA 185

RESULT 7

US-09-919-580-627
; Sequence 627, Application US/09919580
; Patent No. US20020110832A1

GENERAL INFORMATION:

; APPLICANT: Pyle, Ruth
; APPLICANT: Xu, Jiangchun
; APPLICANT: Secrist, Heather
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF COLON CANCER
; FILE REFERENCE: 210121.552
; CURRENT APPLICATION NUMBER: US/09/919,580
; CURRENT FILING DATE: 2001-07-30
; NUMBER OF SEQ ID NOS: 934
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 627
; LENGTH: 434
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-919-580-627

Query Match 67.1%; Score 88.6; DB 10; Length 434;
Best Local Similarity 83.7%; Pred. No. 5.3e-19;
Matches 113; Conservative 0; Mismatches 19; Indels 3; Gaps 1;

QY 1 ATGGCAGACAACTAGACCTGGGAAGAAATTCAGCTTGGATAAGGCCAAGCTGAAGGCC 60
Db 54 ATGGCAGACAACTAGACCTGGGAAGAAATTCAGCTTGGATAAGGCCAAGCTGAAGAAA 113

QY 61 ACAGAGATGC---AGAAGAACACTCTGATGACCAAGAGACCAACAGCAGGAGAGAGTGG 117
Db 114 ACGGAGACGACGAGGAAGAACACCCCTGCCGACCAAGAGACCACTTGCAGCAGGAGAAGCGG 173
QY 118 AGTGAATTTCTCTGA 132
Db 174 AGTGAATTTCTCTAA 188

RESULT 8

US-09-919-580-458
; Sequence 458, Application US/09919580
; Patent No. US20020110832A1

GENERAL INFORMATION:

; APPLICANT: Pyle, Ruth
; APPLICANT: Xu, Jiangchun
; APPLICANT: Secrist, Heather
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF COLON CANCER
; FILE REFERENCE: 210121.552
; CURRENT APPLICATION NUMBER: US/09/919,580
; CURRENT FILING DATE: 2001-07-30
; NUMBER OF SEQ ID NOS: 934
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 458
; LENGTH: 438
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-919-580-458

Query Match 67.1%; Score 88.6; DB 10; Length 438;
Best Local Similarity 83.7%; Pred. No. 5.4e-19;
Matches 113; Conservative 0; Mismatches 19; Indels 3; Gaps 1;

QY 1 ATGGCAGACAACTAGACCTGGGAAGAAATTCAGCTTGGATAAGGCCAAGCTGAAGGCC 60
Db 60 ATGGCAGACAACTAGACCTGGGAAGAAATTCAGCTTGGATAAGGCCAAGCTGAAGAAA 119
QY 61 ACAGAGATGC---AGAAGAACACTCTGATGACCAAGAGACCAACAGCAGGAGAGAGTGG 117
Db 120 ACGGAGACGACGAGGAAGAACACCCCTGCCGACCAAGAGACCACTTGCAGCAGGAGAAGCGG 179
QY 118 AGTGAATTTCTCTGA 132
Db 180 AGTGAATTTCTCTAA 194

RESULT 9

US-09-919-580-583
; Sequence 583, Application US/09919580
; Patent No. US20020110832A1

GENERAL INFORMATION:

; APPLICANT: Pyle, Ruth
; APPLICANT: Xu, Jiangchun
; APPLICANT: Secrist, Heather
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF COLON CANCER
; FILE REFERENCE: 210121.552
; CURRENT APPLICATION NUMBER: US/09/919,580
; CURRENT FILING DATE: 2001-07-30
; NUMBER OF SEQ ID NOS: 934
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 583
; LENGTH: 438
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-919-580-583

Query Match 67.1%; Score 88.6; DB 10; Length 438;
Best Local Similarity 83.7%; Pred. No. 5.4e-19;
Matches 113; Conservative 0; Mismatches 19; Indels 3; Gaps 1;

QY 1 ATGGCACAACACTAGACCTGGAGAAATTCAGCTTGGATAGGCGCAAGCTGAAGGCC 60
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 60 ATGGCACAACACCAGACATGGGGAAATCCAGCTTCGATAGGCCAAGCTGAAGAAA 119
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 61 ACAGAGATGC---AGAAGAACACTCTGTATGACCAAGAGACACAGAGCAGGAGAAATGG 117
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 120 ACGGAGCGCAGGAGAGAACACCCCTGCCGACCAAGAGACCAATTGAGCAGGAGAGCGG 179
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 118 AGTGAATTTCTCTGA 132
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 180 AGTGAATTTCTCTAA 194
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
RESULT 10
US-09-919-580-910
; Sequence 910, Application US/09919580
; Patent No. US20020110832A1
; GENERAL INFORMATION:
; APPLICANT: Pyle, Ruth
; APPLICANT: Xu, Jiangchun
; APPLICANT: Secrist, Heather
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.552
; CURRENT APPLICATION NUMBER: US/09/919,580
; CURRENT FILING DATE: 2001-07-30
; NUMBER OF SEQ ID NOS: 934
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 910
; LENGTH: 439
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 1, 438, 439
; OTHER INFORMATION: n = A,T,C or G
US-09-919-580-910

Query Match 67.1%; Score 88.6; DB 10; Length 439;
Best Local Similarity 83.7%; Pred. No. 5.4e-19;
Matches 113; Conservative 0; Mismatches 19; Indels 3; Gaps 1;
QY 1 ATGGCACAACACTAGACCTGGAGAAATTCAGCTTGGATAGGCGCAAGCTGAAGGCC 60
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 55 ATGGCACAACACCAGACATGGGGAAATCCAGCTTCGATAGGCCAAGCTGAAGAAA 114
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 61 ACAGAGATGC---AGAAGAACACTCTGTATGACCAAGAGACACAGAGCAGGAGAAATGG 117
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 115 ACGGAGCGCAGGAGAGAACACCCCTGCCGACCAAGAGACCAATTGAGCAGGAGAGCGG 174
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 118 AGTGAATTTCTCTGA 132
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 175 AGTGAATTTCTCTAA 189
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 11
US-09-919-580-659
; Sequence 659, Application US/09919580
; Patent No. US20020110832A1
; GENERAL INFORMATION:
; APPLICANT: Pyle, Ruth
; APPLICANT: Xu, Jiangchun
; APPLICANT: Secrist, Heather
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.552
; CURRENT APPLICATION NUMBER: US/09/919,580
; CURRENT FILING DATE: 2001-07-30
; NUMBER OF SEQ ID NOS: 934
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 659
; LENGTH: 443
; TYPE: DNA

; ORGANISM: Homo sapiens
US-09-919-580-659
Query Match 67.1%; Score 88.6; DB 10; Length 443;
Best Local Similarity 83.7%; Pred. No. 5.4e-19;
Matches 113; Conservative 0; Mismatches 19; Indels 3; Gaps 1;
QY 1 ATGGCACAACACTAGACCTGGAGAAATTCAGCTTGGATAGGCGCAAGCTGAAGGCC 60
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 65 ATGGCACAACACCAGACATGGGGAAATCCAGCTTCGATAGGCCAAGCTGAAGAAA 124
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 61 ACAGAGATGC---AGAAGAACACTCTGTATGACCAAGAGACACAGAGCAGGAGAAATGG 117
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 125 ACGGAGCGCAGGAGAGAACACCCCTGCCGACCAAGAGACCAATTGAGCAGGAGAGCGG 184
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 118 AGTGAATTTCTCTGA 132
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 185 AGTGAATTTCTCTAA 199
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
RESULT 12
US-09-919-580-727
; Sequence 727, Application US/09919580
; Patent No. US20020110832A1
; GENERAL INFORMATION:
; APPLICANT: Pyle, Ruth
; APPLICANT: Xu, Jiangchun
; APPLICANT: Secrist, Heather
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.552
; CURRENT APPLICATION NUMBER: US/09/919,580
; CURRENT FILING DATE: 2001-07-30
; NUMBER OF SEQ ID NOS: 934
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 727
; LENGTH: 445
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-919-580-727
Query Match 67.1%; Score 88.6; DB 10; Length 445;
Best Local Similarity 83.7%; Pred. No. 5.4e-19;
Matches 113; Conservative 0; Mismatches 19; Indels 3; Gaps 1;
QY 1 ATGGCACAACACTAGACCTGGAGAAATTCAGCTTGGATAGGCGCAAGCTGAAGGCC 60
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 60 ATGGCACAACACCAGACATGGGGAAATCCAGCTTCGATAGGCCAAGCTGAAGAAA 119
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 61 ACAGAGATGC---AGAAGAACACTCTGTATGACCAAGAGACACAGAGCAGGAGAAATGG 117
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 120 ACGGAGCGCAGGAGAGAACACCCCTGCCGACCAAGAGACCAATTGAGCAGGAGAGCGG 179
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 118 AGTGAATTTCTCTGA 132
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 180 AGTGAATTTCTCTAA 194
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 13
US-09-919-580-891
; Sequence 891, Application US/09919580
; Patent No. US20020110832A1
; GENERAL INFORMATION:
; APPLICANT: Pyle, Ruth
; APPLICANT: Xu, Jiangchun
; APPLICANT: Secrist, Heather
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.552
; CURRENT APPLICATION NUMBER: US/09/919,580
; CURRENT FILING DATE: 2001-07-30
; NUMBER OF SEQ ID NOS: 934
; SOFTWARE: FastSeq for Windows Version 4.0

```

; SEQ ID NO 891
; LENGTH: 446
; TYPE: DNA
; ORGANISM: H
; FEATURE:
; NAME/KEY: m
; LOCATION: 4
; OTHER INFOR:
US-09-919-580-8

```

	Query Match	67.1%	Score 88.6;	DB 10;	Length 446;
	Best Local Similarity	83.7%;	Pred. No. 5.4e-19;		
	Matches 113;	Conservative 0;	Mismatches 19;	Indels 3;	Gaps 1;
QY	1	ATGSCACACAACTAGACCTTGGGAAGAAATTCGCCAGCTTGGATAGGCCCAAGCTCAAGGCC	60		
Db	44	ATGSCACACAAACAGACATATGGGGGAATTCGCCAGCTTCGATAGGCCCAAGCTCAAGAAA	103		
QY	61	ACAGAGATGC---AGAAACAACCTCTGATGACCAAGAGAGACCACAGACGAGGAAGTGG	117		
Db	104	ACGGAGACGCGAGGAGAGAACACCCCTCCCGACCAAGACAGACCATTTGACGAGGAACGG	163		
QY	118	AGTGAATTTCCCTGA	132		
Db	164	AGTGAATTTCCCTAA	178		

RESULT 14

```

US-09-918-995-15768
; Sequence 15768, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID
; TITLE OF INVENTION: FROM VARIOUS CDNA
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,07
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version
; SEQ ID NO 15768
; LENGTH: 449
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-918-995-15768

```

	Query Match	67.1%	Score 88.6;	DB 9;	Length 449;
	Best Local Similarity	83.7%	Pred. No. 5.4e-19;		
	Matches 113;	Conservative 0;	Mismatches 19;	Indels 3;	Gaps 1;
Qy	1	ATGGCACAAACTAGACCTTGGAGAANAATGCCAGCTTGGATATAGGCCCAAGCTGAAGGCC	60		
Db	107	ATGGCACAAACACAGACATGGGGGAATCGCCAGCTTCGATATAGGCCCAAGCTGAAGAAA	166		
Qy	61	ACAGAGATGC---AGAGAACACATCTCATACCACCAAGAGACCCACAGACGACGAGGAATGG	117		
Db	167	ACGGAGACGGCAGGAGAGACACACCCCTGCCGACCAAGAGACCAATTTGACAGAGAGACGG	226		
Qy	118	AGTGAANAATTTCCCTGA	132		
Db	227	AGTGAANAATTTCCCTAA	241		

RESULT 15

; Sequence 553, Application US/09919580
 ; Patent No. US20020110832A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Pyle, Ruth
 ; APPLICANT: Xu, Jiangchun
 ; APPLICANT: Secrist, Heather

```

: TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
:
: TITLE OF INVENTION: AND DIAGNOSIS OF COLON CANCER
:
: FILE REFERENCE: 210121.552
:
: CURRENT APPLICATION NUMBER: US/09/919,580
:
: CURRENT FILING DATE: 2001-07-30
:
: NUMBER OF SEQ ID NOS: 934
:
: SOFTWARE: FastSEQ for Windows Version 4.0
:
: SEQ ID NO 553
:
: LENGTH: 451
:
: TYPE: DNA
:
: ORGANISM: Homo sapiens
:
: US-09-919-580-553

```

Query Match	67.1%	Score 98.6	DB 10	Length 451
Best Local Similarity	83.7%	Pred. No. 5.4e-19		
Matches 113	Conservative 0	Mismatches 19	Indels 3	Gaps 1
Qy	1	ATGSCACACAAACTPAGACCTTGGAGAAGAAATTCGCCAGCTTGGATATAGGCCCAAGCTGAAGGCC	60	
Db	73	ATGCGACAGACAAACACAGACATCTGGGGGAATTCGCCAGCTTCGATATAGGCCCAAGCTGAGAAA	132	
Qy	61	ACAGAGATGC---AGAAGACACACTCTGATGACCAGAAAGAGACCCACAGACGAGAGAAAGTGG	117	
Db	133	ACGGAGACGAGGAGAGAACACACCTGCGGACCAAGAGACCATTTAGCAGCAGGAAGCGG	192	
Qy	118	AGTGAATTTTCCTGA	132	
Db	193	AGTGAATTTTCCTAA	207	

Search completed: June 2, 2003, 15:41:27
Job time : 107 secs

THIS PAGE BLANK (10PT0)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run On: June 2, 2003, 14:12:12 ; Search time 1421 Seconds

(without alignments)
1504.436 Million cell updates/sec

Title: US-09-915-178-1

Perfect score: 132

Sequence: 1 atggcacacaaactagacct.....agtggagtgaatttcctga 132

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

```
EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estinu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_nam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rod:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result	No.	Score	Query Match	Length	DB ID	Description
c	1	90.4	68.5	210	10 AW240381	AW240381 up31e02.y
	2	90.2	68.3	299	14 BM855208	BM855208 K-EST0138
c	3	90.2	68.3	475	9 AA525997	AA525997 nt58b07.s
c	4	89.2	67.6	403	9 AI708863	AI708863 as27103.x
	5	89.2	67.6	867	12 BG831864	BG831864 602765514
	6	88.6	67.1	174	14 T27818	T27818 EST17484 Hu

7	88.6	67.1	182	9	AA316715	AA316715 EST188545
8	88.6	67.1	207	9	AA312784	AA312784 EST183467
9	88.6	67.1	213	14	F26549	F26549 HSPD14071 H
10	88.6	67.1	215	10	AW369977	AW369977 IL0-BT023
11	88.6	67.1	217	14	D53771	D53771 HOM119F07B
12	88.6	67.1	227	14	T60895	T60895 YB72F05.r1
13	88.6	67.1	239	9	AA340512	AA340512 EST45729
14	88.6	67.1	239	9	AA366141	AA366141 EST77049
15	88.6	67.1	240	14	T25985	T25985 ATH467 HTCD
16	88.6	67.1	241	9	AA302528	AA302528 EST15689
17	88.6	67.1	247	9	AA362157	AA362157 EST17636
18	88.6	67.1	251	9	AA301772	AA301772 EST14827
19	88.6	67.1	254	14	N85477	N85477 J3657F Huma
20	88.6	67.1	257	9	AA354418	AA354418 EST62702
21	88.6	67.1	257	10	AW800885	AW800885 MR3-UM006
22	88.6	67.1	258	9	AA339727	AA339727 EST44841
23	88.6	67.1	258	13	BG994832	BG994832 PM0-HT116
24	88.6	67.1	260	9	AA365696	AA365696 EST76526
25	88.6	67.1	260	10	AW998455	AW998455 PM2-BN006
26	88.6	67.1	264	12	BF378584	BF378584 CM0-UM004
27	88.6	67.1	270	9	AA304229	AA304229 EST16960
28	88.6	67.1	271	9	AA354636	AA354636 EST62927
29	88.6	67.1	273	10	AW797766	AW797766 CM2-UM004
30	88.6	67.1	273	10	AW800972	AW800972 MR3-UM006
31	88.6	67.1	274	9	AA308404	AA308404 EST175234
32	88.6	67.1	274	10	AW797701	AW797701 CM0-UM004
33	88.6	67.1	276	9	AA365522	AA365522 EST76310
34	88.6	67.1	276	14	BM855353	BM855353 K-EST0138
35	88.6	67.1	279	9	AA362030	AA362030 EST71374
36	88.6	67.1	280	10	AW797700	AW797700 CM0-UM004
37	88.6	67.1	282	9	AA301934	AA301934 EST15013
38	88.6	67.1	285	12	BE714668	BE714668 PM2-HT072
39	88.6	67.1	288	9	AA338821	AA338821 EST43818
40	88.6	67.1	289	9	AA374209	AA374209 EST86338
41	88.6	67.1	290	9	AA070672	AA070672 2m53g03.r
42	88.6	67.1	290	9	AA380802	AA380802 EST93797
43	88.6	67.1	292	9	AA304555	AA304555 EST17331
44	88.6	67.1	293	9	AA301999	AA301999 EST15069
45	88.6	67.1	295	9	AA384610	AA384610 EST98393

ALIGNMENTS

RESULT 1
AW240381/c
LOCUS AW240381 210 bp mRNA linear EST 14-DEC-1999
DEFINITION up31e02.y1 NCI-CGAP_Mam2 Mus musculus cDNA clone IMAGE:2655962 5' similar to gb:SS4005 THYMOSIN BETA-10 (HUMAN);, mRNA sequence.
ACCESSION AW240381
VERSION AW240381.1 GI:6574067
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
AUTHORS 1 (bases 1 to 210)
TITLE NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
JOURNAL National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Unpublished (1997)
COMMENT Tumor Gene Index
Email: cgaps-r@mail.nih.gov
Contact: Robert Strausberg, Ph.D.
Tissue Procurement: Gilbert Smith, Ph.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
Clone Distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: www-bio.llnl.gov/bbrp/image/image.html
MGI:1034110
Seq primer: -40RP from Gibco

High quality sequence stop: 197.
Location/Qualifiers

source
1. .210
/organism="Mus musculus"
/strain="FVB/N-3"
/db_xref="taxon:10090"
/clone_lib="NCI_CGAP_Mam2"
/tissue_type="tumor, biopsy sample"
/dev_stage="5 months"
/lab_host="DH10B"
/note="Organ: mammary; Vector: pCMV-SPORT6; Site_1: SalI;
Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT.
Library constructed by Life Technologies. Investigator
providing samples: Gilbert Smith, NIH"

BASE COUNT 26 a 60 c 46 g 78 t
ORIGIN
Query Match 68.5%; Score 90.4; DB 10; Length 210;
Best Local Similarity 80.3%; Pred. No. 2.6e-15;
Matches 106; Conservative 0; Mismatches 26; Indels 0; Gaps 0;
QY 1 ATGGCAGACAACTAGACCTGGAAGAAATTCAGCTGGATAAGGCCAAGCTGAAGGCC 60
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 154 ATGGCAGACAACTAGACCTGGAAGAAATTCAGCTGGATAAGGCCAAGCTGAAGGCC 95
QY 61 ACAGAGATGCGAGAGAACACTCTGTATGACCAAGAGACACAGAGAGAGAGAGAGTGGAGT 120
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 94 ACCGAGATGCGAGAGAACACTCTGTATGACCAAGAGACACAGAGAGAGAGAGTGGAGT 120
QY 121 GAAATTTCTCTGA 132
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 34 GAAATTTCTCTGA 23

RESULT 2
BM855208
LOCUS
DEFINITION K-EST0138016 S14K402 Homo sapiens cDNA clone S14K402-46-E08 5',
mRNA sequence.
ACCESSION BM855208
VERSION BM855208.1 GI:19211607
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 299)
Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
Kim,Y.S.
21C Frontier Korean EST Project 2001
Unpublished (2002)
Contact: Kim YS
Genome Research Center
Korea Research Institute of Bioscience & Biotechnology
52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
Tel: +82-42-860-4470
Fax: +82-42-860-4409
Email: yongsung@mail.kribb.re.kr
Plate: 46 Row: E Column: 08
High quality sequence stop: 299.
Location/Qualifiers

FEATURES
source
1. .299
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="S14K402-46-E08"
/clone_lib="S14K402"
/cell_line="K402"
/lab_host="Top10F"
/note="Organ: Stomach; Vector: pTZ18Rpl; Site_1: EcoRI;
Site_2: NotI; The poly (A)+ RNA was dephosphorylated with
bacterial alkaline phosphatase (BAP) and then decapped

with tobacco acid pyrophosphatase (TAP). The decapped
intact mRNA was ligated with DNA-RNA linker including EcoR
I site by treatment of T4 RNA ligase and the first strand
cDNA was synthesized from oligo dT-selected mRNA by
priming with dT-tailed vector. The dT-tailed vector was
adjusted to have about 60nt. The cDNA vector was
circularized with E. coli DNA ligase after digestion of
EcoRI which site is also included in vector. An RNA strand
converted to a DNA strand by Okayama-Berg method. The
obtained cDNA vectors were used for transposition of
competent cells E. coli Top10F by electroporation method.
The cDNA libraries constructed by this method are
full-length enriched cDNA library."

BASE COUNT 100 a 67 c 82 g 50 t
ORIGIN
Query Match 68.3%; Score 90.2; DB 14; Length 299;
Best Local Similarity 84.4%; Pred. No. 3.5e-15;
Matches 114; Conservative 0; Mismatches 18; Indels 3; Gaps 1;
QY 1 ATGGCAGACAACTAGACCTGGAAGAAATTCAGCTGGATAAGGCCAAGCTGAAGGCC 60
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 101 ATGGCAGACAACTAGACCTGGAAGAAATTCAGCTGGATAAGGCCAAGCTGAAGGCC 160
QY 61 ACAGAGATGCG---AGAGAGAACACTCTGTATGACCAAGAGACACAGAGAGAGAGTGG 117
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 161 ACAGAGATGCGAGAGAACACTCTGTATGACCAAGAGACACAGAGAGAGAGTGG 117
QY 118 AGTGAATTTCTCTGA 132
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 221 AGTGAATTTCTCTGA 235

RESULT 3
AA525997/c
LOCUS
DEFINITION AA525997 n158b07.s1 NCI_CGAP_Ov2 475 bp mRNA linear EST 05-AUG-1997
similar to gb:S54005 THYMOSIN BETA-10 (HUMAN);, mRNA sequence.
ACCESSION AA525997
VERSION AA525997.1 GI:2268066
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 475)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Christopher A. Moskaluk, M.D., Michael R.
Emmert-Buck, M.D., Ph.D.
CDNA Library Preparation: David B. Krizman, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Insert Length: 561 Std Error: 0.00
Seq primer: -40ml3 fwd. Et from Amersham
High quality sequence stop: 212.
Location/Qualifiers

FEATURES
source
1. .475
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:981013"
/clone_lib="NCI_CGAP_Ov2"
/sex="female"
/tissue_type="ovary"
/lab_host="DH10B"
/note="Vector: pAMP10; mRNA made from invasive ovarian

Query Match 67.1%; Score 88.6; DB 9; Length 207;
Best Local Similarity 83.7%; Pred. No. 8.4e-15;

1	ATGGCACACAACTAGACCTGGAGAAATTCGCCAGCTTGGATGAAGCCCAAGCTGAAGGCC	60
QY		
18	ATGGCGAGACAAACACAGACATCGGGGAAATCGCCAGCTTCGATGAAGCCCAAGCTGAAGAAA	77
Db		
61	ACAGAGATGC---AAGAGAAACACTCTGTATGACCAAGAGACACCAAGAGAGAGAAAGTGG	117
QY		
78	ACGGGAGACGACGAGGAGAAACACCTTCGCCACCAAGAGACCAATTTCAGCAGGAGAGACGG	137
Db		

97044478
Contact: Wilson RK
Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: estewatson.wustl.edu

Insert Size: 18

High quality sequence stops: 191 Source: IMAGE Consortium, LLNL This clone is available royalty-free through LLNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information. Putative full length read.

Seq primer: M13R21

High quality sequence stop: 191.

Location/Qualifiers

FEATURES

source

1. .227
 /organism="Homo sapiens"
 /db_xref="GDB:506466"
 /db_xref="taxon:9606"
 /clone="IMAGE:76737"
 /clone_lib="Stratagene ovary (#937217)"
 /sex="female"
 /dev_stage="49 year old"
 /lab_host="SOLR cells (kanamycin resistant)"
 /note="organ: ovary; Vector: Bluescript SK; Site_1: EcoRI; Site_2: XhoI; Cloned unidirectionally. Primer: Oligo df. Total ovary tissue, normal, caucasian. Average insert size: 0.8 kb; Uni-ZAP XR Vector; -5' adaptor sequence: 5' GAATTCGGCAGCAG 3' -3' adaptor sequence: 5' CTCGAGTGTCTTTTCTTTT 3'"
 BASE COUNT 70 a 51 c 64 g 39 t 3 others
 ORIGIN

Query Match 67.1%; Score 88.6; DB 14; Length 227;

Best Local Similarity 83.7%; Pred. No. 8.7e-15;

Matches 113; Conservative 0; Mismatches 19; Indels 3; Gaps 1;

QY 1 ATGSCACACAACTGACCTGGAAGAAATGCCAGCTTGATGAAGCCAGCTGAGGCC 60

Db 60 ATGGCAGACAAACCCAGACATCGGGAAATCGCCAGCTTCGATAAGCCAGCTGAAGAA 119

QY 61 ACAGAGATGC---AGAAGAACACTCTGATGACCAAGAGACACAGCAGGAGAGAGTGG 117

Db 120 ACGGAGACGACAGAGAGACACCTCGCCGACCAAGAGACCAATTGAGCAGGAGAGCGG 179

QY 118 AGTGAATTTCTCTGA 132

Db 180 AGTGAATTTCTCTAA 194

RESULT 13

AA340512

LOCUS

DEFINITION EST45729 Fetal kidney III Homo sapiens cDNA 3' end similar to

similar to thymosin beta-10 (GB:S54005), mRNA sequence.

AA340512.1 GI:1992750

ACCESSION

VERSION

KEYWORDS

SOURCE

human.

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 239)

AUTHORS

Adams, M.D., Kerlavage, A.R., Fleischmann, R.D., Fuldner, R.A., Bult
 , C.J., Lee, N.H., Kirkness, E.F., Weinstock, K.G., Gocayne, J.D., White
 , O., Sutton, G., Blake, J.A., Brandon, R.C., Man-Wai, C., Clayton, R.A.,
 Cline, T.R., Cotton, M.D., Earle-Hughes, J., Fine, L.D., Fitzgerald
 , L.M., Fitzhugh, W.M., Fritchman, J.L., Geoghagen, N.S., Glodek, A.,
 Kelley, J.C., Liu, L.-I., Marmaros, S.M., Merrick, J.M.,
 Moreno-Palauques, R.F., McDonald, L.A., Nguyen, D.T., Pelligrino, S.M.,
 Phillips, C.A., Ryder, S.E., Scott, J.L., Saudek, D.M., Shirley, R.,
 Small, K.V., Spriggs, T.A., Utterback, T.R., Weidman, J.F., Li, Y.,
 Dimke, D., Feng, D.-F., Ferrie, A., Coleman, T.A., Hastings, G.A., He, W.W.,
 Hu, J.S., Greene, J.M., Gruber, J., Hudson, P., Kim, A.K., Kozak, D.L.,

Kunsch, C., Hunglun, J., Li, H., Weissner, P.S., Olsen, H., Raymond, L.,
 Wei, Y.F., Wing, J., Xu, C., Yu, G.L., Ruben, S.M., Dillion, P.J., Fannon
 , M.R., Rosen, C.A., Haseltine, W.A., Fields, C., Fraser, C.M. and
 Venter, J.C.

Initial assessment of human gene diversity and expression patterns
 based upon 83 million nucleotides of cDNA sequence
 Nature 377 (6547 Suppl), 3-174 (1995)

96026280

Contact: Kerlavage, AR

Bioinformatics

The Institute for Genomic Research

9712 Medical Center Drive, Rockville, MD 20850 USA

Tel: 3018699056

Fax: 3018699423

Email: arkerlav@tigr.org

For clone availability, additional sequence and expression

information related to this EST, please check the TIGR Human Gene

Index (<http://www.tigr.org/tdb/hgi/hgi.html>)

Seq primer: M13 Reverse.

Location/Qualifiers

1. .239
 /organism="Homo sapiens"
 /db_xref="ATCC (Inhost):142154"
 /db_xref="taxon:9606"
 /clone_lib="Fetal kidney III"
 /sex="female"
 /dev_stage="fetus, 24 wks"
 /note="Organ: kidney; Vector: pBluescript KS-; Site_1:
 XhoI; Site_2: EcoRI"

BASE COUNT 72 a 56 g 39 t 3 others

ORIGIN

Query Match 67.1%; Score 88.6; DB 9; Length 239;

Best Local Similarity 83.7%; Pred. No. 8.9e-15;

Matches 113; Conservative 0; Mismatches 19; Indels 3; Gaps 1;

QY 1 ATGSCACACAACTGACCTGGAAGAAATGCCAGCTTGATGAAGCCAGCTGAGGCC 60

Db 53 ATGGCAGACAAACCCAGACATCGGGAAATCGCCAGCTTCGATAAGCCAGCTGAAGAA 112

QY 61 ACAGAGATGC---AGAAGAACACTCTGATGACCAAGAGACACACAGCAGGAGAGTGG 117

Db 113 ACGGAGACGACAGGAGAGACACCTCGCCGACCAAGAGACCAATTGAGCAGGAGAGCGG 172

QY 118 AGTGAATTTCTCTGA 132

Db 173 AGTGAATTTCTCTAA 187

RESULT 14

AA366141

LOCUS

DEFINITION EST77049 Pancreas tumor III Homo sapiens cDNA 5' end similar to

similar to thymosin beta-10 (GB:S54005), mRNA sequence.

AA366141

ACCESSION

VERSION

KEYWORDS

SOURCE

human.

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 239)

AUTHORS

Adams, M.D., Kerlavage, A.R., Fleischmann, R.D., Fuldner, R.A., Bult
 , C.J., Lee, N.H., Kirkness, E.F., Weinstock, K.G., Gocayne, J.D., White
 , O., Sutton, G., Blake, J.A., Brandon, R.C., Man-Wai, C., Clayton, R.A.,
 Cline, T.R., Cotton, M.D., Earle-Hughes, J., Fine, L.D., Fitzgerald
 , L.M., Fitzhugh, W.M., Fritchman, J.L., Geoghagen, N.S., Glodek, A.,
 Gnehm, C.L., Hanna, M.C., Hedblom, E., Hinkle, P.S., Jr., Kelley, J.M.,
 Kelley, J.C., Liu, L.-I., Marmaros, S.M., Merrick, J.M.,
 Moreno-Palauques, R.F., McDonald, L.A., Nguyen, D.T., Pelligrino, S.M.,
 Phillips, C.A., Ryder, S.E., Scott, J.L., Saudek, D.M., Shirley, R.,
 Small, K.V., Spriggs, T.A., Utterback, T.R., Weidman, J.F., Li, Y.,
 Bednarik, D.P., Cao, L., Cepeda, M.A., Coleman, T.A., Collins, E.J.,
 Bednarik, D.P., Cao, L., Cepeda, M.A., Coleman, T.A., Collins, E.J.,

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 2, 2003, 13:25:22 ; Search time 1074 Seconds
(without alignments)
3576.881 Million cell updates/sec

Title: US-09-915-178-1
Perfect score: 132
Sequence: 1 atggcacacaaactagacct.....agtggagtgaattctctga 132

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*

- 1: gb_ba.*
- 2: gb_hgt.*
- 3: gb_in.*
- 4: gb_om.*
- 5: gb_ov.*
- 6: gb_pat.*
- 7: gb_ph.*
- 8: gb_pl.*
- 9: gb_pr.*
- 10: gb_ro.*
- 11: gb_sts.*
- 12: gb_sv.*
- 13: gb_un.*
- 14: gb_vi.*
- 15: em_ba.*
- 16: em_fun.*
- 17: em_hum.*
- 18: em_in.*
- 19: em_mu.*
- 20: em_om.*
- 21: em_or.*
- 22: em_ov.*
- 23: em_pat.*
- 24: em_ph.*
- 25: em_pl.*
- 26: em_ro.*
- 27: em_sts.*
- 28: em_un.*
- 29: em_vi.*
- 30: em_htg_hum.*
- 31: em_htg_inv.*
- 32: em_htg_other.*
- 33: em_htg_mus.*
- 34: em_htg_pln.*
- 35: em_htg_rod.*
- 36: em_htg_mam.*
- 37: em_htg_vrt.*
- 38: em_sy.*
- 39: em_htgo_hum.*
- 40: em_htgo_mus.*
- 41: em_htgo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	132	100.0	147	6	AX399230	Sequence
2	132	100.0	147	6	AX399232	Sequence
3	132	100.0	147	6	AX399234	Sequence
4	132	100.0	170536	9	HS18K17	Sequence
5	90.4	68.5	154265	2	AC127915	Homo sapi
6	90.4	68.5	158286	2	AC111338	Rattus no
7	90.4	68.5	194568	2	AC129614	Rattus no
8	88.8	67.3	183583	2	AC131114	Rattus no
9	88.8	67.3	186744	2	AC131561	Mus muscu
10	88.8	67.3	336512	2	AC126029	Rattus no
11	88.6	67.1	174	11	G06175	Mus muscu
12	88.6	67.1	243	6	AX381579	Sequence
13	88.6	67.1	353	6	AX381286	Sequence
14	88.6	67.1	395	6	AX381269	Sequence
15	88.6	67.1	400	9	HUMTHMBX	Human thymo
16	88.6	67.1	428	6	AX381565	Sequence
17	88.6	67.1	429	6	AX381695	Sequence
18	88.6	67.1	434	6	AX381689	Sequence
19	88.6	67.1	438	6	AX381520	Sequence
20	88.6	67.1	438	6	AX381645	Sequence
21	88.6	67.1	439	6	AX381972	Sequence
22	88.6	67.1	439	9	HUMTHYB10	Human thymo
23	88.6	67.1	443	6	AX381721	Sequence
24	88.6	67.1	445	6	AX381789	Sequence
25	88.6	67.1	446	6	AX381953	Sequence
26	88.6	67.1	451	6	AX381615	Sequence
27	88.6	67.1	452	6	AX381410	Sequence
28	88.6	67.1	453	6	AX381336	Sequence
29	88.6	67.1	453	6	AX410373	Sequence
30	88.6	67.1	453	9	S54005	thymosin be
31	88.6	67.1	455	6	AX381335	Sequence
32	88.6	67.1	456	6	AX381281	Sequence
33	88.6	67.1	465	6	AX381118	Sequence
34	88.6	67.1	491	6	AX107107	Sequence
35	88.6	67.1	500	9	BC016731	Homo sapi
36	88.6	67.1	511	9	BC016025	Homo sapi
37	88.6	67.1	517	6	AX381936	Sequence
38	88.6	67.1	693	6	AX381860	Sequence
39	88	66.7	430	6	AX119979	Sequence
40	88	66.7	107467	9	HS95883	Homo sapien
41	85.6	64.8	460	6	ARI44218	Sequence
42	85.6	64.8	460	6	ARI76405	Sequence
43	85.6	64.8	460	6	AX365690	Sequence
44	85.6	64.8	155930	2	AC109106	Rattus no
45	85.4	64.7	465	4	AF506973	Equus cab

ALIGNMENTS

RESULT 1	AX399230	Sequence 3 from Patent WO0190155.	147 bp	DNA	linear	PAT 28-MAY-2002
LOCUS	AX399230	Sequence 3 from Patent WO0190155.				
DEFINITION	AX399230					
ACCESSION	AX399230					
VERSION	AX399230.1	GI:21261549				
KEYWORDS	human.					
SOURCE	human.					
ORGANISM	Homo sapiens					
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;					
	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.					
REFERENCE	1					
AUTHORS	Spytek, K.A., Majumder, K., Tchernev, V.T., Mishra, V., Padigar, M., Spaderna, S.K., Shenoy, S., Rastelli, L., Li, L., Taupier, R.J. and Gangolli, E.					

TITLE Novel proteins and nucleic acids encoding same
JOURNAL Patent: WO 0190155-A 3 29-NOV-2001;
Curagen Corporation (US)
FEATURES Location/Qualifiers
source 1..147

BASE COUNT 56 a 30 c 41 g 20 t
ORIGIN /organism="Homo sapiens"
/db_xref="taxon:9606"

Query Match 100.0%; Score 132; DB 6; Length 147;
Best Local Similarity 100.0%; Pred. No. 9.1e-27;
Matches 132; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCACAACTAGACCTGGAGAAATTCAGCTTGGATAGGCCAAGCTGAAGGCC 60
|||||
Db 6 ATGGCACAACTAGACCTGGAGAAATTCAGCTTGGATAGGCCAAGCTGAAGGCC 65
|||||
QY 61 ACAGAGATGCAGAGACACTCTGTGATGACCAAGAGACACAGAGAGAGAGTGGAGT 120
|||||
Db 66 ACAGAGATGCAGAGACACTCTGTGATGACCAAGAGACACAGAGAGAGAGTGGAGT 120
|||||
QY 121 GAAATTTCTCTGA 132
|||||
Db 126 GAAATTTCTCTGA 137

RESULT 2
AX399232
LOCUS Sequence 5 from Patent WO0190155. 147 bp DNA linear PAT 28-MAY-2002
DEFINITION
ACCESSION
VERSION
AX399232.1 GI:21261550
KEYWORDS
SOURCE human.

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1

REFERENCE
AUTHORS Spytek, K.A., Majumder, K., Tchernev, V.T., Mishra, V., Padigar, M.,
Spaderna, S.K., Shenoy, S., Rastelli, L., Li, L., Taupier, R.J., and
Gangolli, E.

TITLE Novel proteins and nucleic acids encoding same
JOURNAL Patent: WO 0190155-A 5 29-NOV-2001;
Curagen Corporation (US)

FEATURES Location/Qualifiers
source 1..147

BASE COUNT 56 a 30 c 41 g 20 t
ORIGIN /organism="Homo sapiens"
/db_xref="taxon:9606"

Query Match 100.0%; Score 132; DB 6; Length 147;
Best Local Similarity 100.0%; Pred. No. 9.1e-27;
Matches 132; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCACAACTAGACCTGGAGAAATTCAGCTTGGATAGGCCAAGCTGAAGGCC 60
|||||
Db 6 ATGGCACAACTAGACCTGGAGAAATTCAGCTTGGATAGGCCAAGCTGAAGGCC 65
|||||
QY 61 ACAGAGATGCAGAGACACTCTGTGATGACCAAGAGACACAGAGAGAGAGTGGAGT 120
|||||
Db 66 ACAGAGATGCAGAGACACTCTGTGATGACCAAGAGACACAGAGAGAGAGTGGAGT 120
|||||
QY 121 GAAATTTCTCTGA 132
|||||
Db 126 GAAATTTCTCTGA 137

RESULT 3
AX399234
LOCUS Sequence 7 from Patent WO0190155. 147 bp DNA linear PAT 03-JUN-2002
DEFINITION

ACCESSION AX399234
VERSION AX399234.1 GI:21261551
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1

REFERENCE

AUTHORS

TITLE Novel proteins and nucleic acids encoding same
JOURNAL Patent: WO 0190155-A 7 29-NOV-2001;
Curagen Corporation (US)

FEATURES Location/Qualifiers
source 1..147

BASE COUNT 56 a 30 c 41 g 20 t
ORIGIN /organism="Homo sapiens"
/db_xref="taxon:9606"

Query Match 100.0%; Score 132; DB 6; Length 147;
Best Local Similarity 100.0%; Pred. No. 9.1e-27;
Matches 132; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCACAACTAGACCTGGAGAAATTCAGCTTGGATAGGCCAAGCTGAAGGCC 60
|||||
Db 6 ATGGCACAACTAGACCTGGAGAAATTCAGCTTGGATAGGCCAAGCTGAAGGCC 65
|||||
QY 61 ACAGAGATGCAGAGACACTCTGTGATGACCAAGAGACACAGAGAGAGTGGAGT 120
|||||
Db 66 ACAGAGATGCAGAGACACTCTGTGATGACCAAGAGACACAGAGAGAGTGGAGT 120
|||||
QY 121 GAAATTTCTCTGA 132
|||||
Db 126 GAAATTTCTCTGA 137

RESULT 4
HS518K17
LOCUS Homo sapiens chromosome 9 BAC RP11-518K17, complete sequence.
DEFINITION
ACCESSION AL513423 AL353720
VERSION AL513423.2 GI:18539155
KEYWORDS HTG.
SOURCE human.

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Submitted (07-FEB-2001) GBF, Dept. of Genome Analysis, Mascheroder
Weg 1, D-38124 Braunschweig, Germany, E-mail: info.genome@gbf.de
On Feb 5, 2002 this sequence version replaced gi:12718191.
All annotations in this database entry are developed by
computational tools. It is therefore not explicitly noted in the
feature lines that evidence is not experimental.
Mapping was performed at The Sanger Centre
(cf. <http://www.sanger.ac.uk/HGP/Chr9>)
Mapping information is available via
<http://webace.sanger.ac.uk/cgi-bin/display?db=acedb&grep=518K17>

Center: GBF, Braunschweig
Center code: GBF
Web site: <http://genome.gbf.de/>

```

Contact: info.genome@bf.de
----- Project Information
Center project name:
Center clone name: BA518K17
----- Summary Statistics
Sequencing vector: ##;
Chemistry: Dye-terminator-BigDye: 33% of reads
Chemistry: Dye-terminator-amersham: 55% of reads
Chemistry: Dye-primer-amersham: 12% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 0 bases at least Q40
Consensus quality: 0 bases at least Q30
Consensus quality: 0 bases at least Q20
Estimated insert size: ##; agarose-fp estimation
Estimated insert size: 170536; sum-of-contigs estimation
-----
PROGRAMS AND PARAMETERS USED FOR ANNOTATION:
+++++
+ Analysis and annotation were performed with the automatic +
+ 'first-pass' annotation and submission tool +
+ 'AnnoMitter' (Hornischer & Bloecker). +
+ Programs used by 'AnnoMitter': +
+++++
> GeneFinder (Green), Vers. 084
+ Organism: human
+ GenScan (Burge & Karlin), Vers. 1.0
+ Used matrix: vertebrate; Minimum score: 0
+ Grail (Xu et al.), Vers. 1.3
+ Organism: human
+ Mzef (Zhang)
+ Prior probability: 0.04; Overlapping number: 0 > Xpound (Thomas
  & Skolnick)
+ BLASTN 2.0.14 (Altschul et al.)
+ Database(s): * RepBase: ALU (human), released 22-DEC-1995
+ * RepBase: THR (human), released 22-DEC-1995
+ * RepBase: MIR (primate), released 22-DEC-1995
+ RepBase: MER (primate), released 22-DEC-1995
+ RepBase: MIR2 (primate), released 22-DEC-1995
+ RepBase: TRE (primate), released 22-DEC-1995
+ Minimum identity: 70 %; Minimum score: 60;
> 'ESTs': BLASTN 2.0.14 (Altschul et al.)
+ Database(s): * emb1 (EST, human), released -DEC-
  (EST), Vers. 67+ (01-JAN-1970) . Using sequence with masked
  repeats
+ Minimum score: 60; Minimum identity: 70 %;
> 'Tandem Repeats': GDE 2.2 option 'tandem'
+ Minimum length 2 bp; Maximum length 20 bp; Score threshold 20 .
+ Treat N's as mismatches? YES; Allow uniform consensi? NO >
> 'Inverted Repeats': GDE 2.2 option 'inverted'
> 'Micro Satellites': GDE 2.2 option 'sputnik' (Abajian) > 'Cpg
  Islands': GDE 2.2 option 'cpg'
+ Cpg island region size 100 bp;
+ Minimum GC contents 50 %; Observed/Expected 0.6 > 'STS Scan':
e-PCR (Schuler)
+ Margin: 50; Number of mismatches allowed: 0; Word size: 7 .
STS database: 'dbSTS markers'
> 'tRNA Scan': tRNAscan-SE (Lowe & Eddy), Vers. 1.11.
Location/Qualifiers
1..170536
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="9"
/clone="RP11-518K17"
/clone_lib="RPCI-11.2"
1..170536
/note="assembly-fragment-clone_end:T7-vector_side:left
assembly-fragment-clone_end:SP6-vector_side:right"
129..177
exon
/note="XPOUND prediction, score = 0.468"
233..247
satellite

repeat_region
/note="AT repeat"
368..415
/note="95% identity: matches 177..224 of consensus"
/rpt_family="L1"
satellite
1012..1023
/note="TA repeat"
repeat_region
1390..1427
/note="IR1, 76% complementary to IR1' (2545..2582)"
/rpt_type=INVERTED
2114..2233
/note="homology = 66.7%, counts = 6"
/rpt_family="atattattttttattattt repeat"
/rpt_type=TANDEM
complement(2225..6898)
/note="94% identity: matches 1344..6024 of consensus"
/rpt_family="L1"
2545..2582
/note="IR1', 76% complementary to IR1 (1390..1427)"
/rpt_type=INVERTED
complement(3019..3294)
/note="GRAIL, score = 59%, comment = good"
complement(3565..4554)
/note="GRAIL, score = 82%, comment = excellent"
complement(3565..4479)
/note="GENSCAN prediction, score = 20.66"
complement(3802..3953)
/note="MzEF prediction, score = 0.634"
4560..4665
/note="GRAIL, score = 77%, comment = excellent shadow"
complement(6316..7104)
/note="GRAIL, score = 67%, comment = good"
complement(6316..7062)
/note="GENSCAN prediction, score = 35.33"
complement(6975..7280)
/note="92% identity: matches 958..1267 of consensus"
/rpt_family="L1"
complement(7134..7262)
/note="GRAIL, score = 69%, comment = good"
complement(7341..7593)
/note="88% identity: matches 1..253 of consensus"
/rpt_family="L1"
7479..7596
/note="XPOUND prediction, score = 0.349"
complement(7586..8358)
/note="93% identity: matches 7..776 of consensus"
/rpt_family="L1"
7953..8079
/note="GRAIL, score = 51%, comment = good shadow"
7964..8094
/note="XPOUND prediction, score = 0.547"
complement(8287..8340)
/note="XPOUND prediction, score = 0.781"
8442..12059
/note="92% identity: matches 2388..6006 of consensus"
/rpt_family="L1"
9004..9080
/note="homology = 70.1%, counts = 7"
/rpt_family="aaaaagcaaga repeat"
/rpt_type=TANDEM
10318..10457
/note="GRAIL, score = 70%, comment = good shadow"
complement(10763..10799)
/note="GRAIL, score = 77%, comment = excellent"
11782..11861
/note="GRAIL, score = 47%, comment = marginal shadow"
complement(12003..12016)
/note="XPOUND prediction, score = 0.218"
12893..12944
/note="homology = 82.7%, counts = 26"
/rpt_family="tg repeat"
/rpt_type=TANDEM
12895..12926
/note="TG repeat"
satellite

```

FEATURES
source

misc_feature

exon

satellite

```

satellite 12927..12937
/note="CG repeat"
complement(12989..13135)
/note="GRAIL, score = 90%, comment = excellent
MZF prediction, score = 0.97"
complement(13577..13664)
/note="GRAIL, score = 66%, comment = good"
complement(14120..14818)
/note="82% identity: matches 399..1100 of consensus"
/rpt_family="L1"
14916..15001
/note="GRAIL, score = 91%, comment = excellent shadow"
complement(15145..15418)
/note="81% identity: matches 312..587 of consensus"
/rpt_family="L1"
15147..15418
/note="84% identity: matches 11..281 of consensus"
/rpt_family="ALU"
15421..15435
/note="AAAT repeat"
complement(15473..15595)

Query Match 100.0%; Score 132; DB 9; Length 170536;
Best Local Similarity 100.0%; Pred. No. 1.5e-26;
Matches 132; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCACAAACTAGACCTGGGAAGAAATTCAGCTGATAGCCAGCTGAAGGCC 60
|||||
Db 111695 ATGGCACAAACTAGACCTGGGAAGAAATTCAGCTGATAGCCAGCTGAAGGCC 111754
|||||

QY 61 ACAGAGATCAGAGAACACTCTGATGACCAAGACACACAGAGAGAGTGGAGT 120
|||||
Db 111755 ACAGAGATCAGAGAACACTCTGATGACCAAGACACACAGAGAGAGTGGAGT 111814
|||||

QY 121 GAAATTCCTGA 132
|||||
Db 111815 GAAATTCCTGA 111826

RESULT 5
AC127915/c
LOCUS AC127915.1 GI:21908439
DEFINITION Rattus norvegicus clone CH230-236P17, *** SEQUENCING IN PROGRESS
***, 61 unordered pieces.
AC127915
VERSION AC127915.1
KEYWORDS HTGS, PHASE1.
SOURCE Rattus norvegicus.
ORGANISM Rattus norvegicus.

REFERENCE
AUTHORS Muzny, D.M., Adams, C., Adio-Oduola, B., Ali-Osman, F.R., Allen, C.,
Alsbrooks, S.L., Anaratinge, H.C., Are, J.R., Ayale, M., Banks, T.,
Barbaria, J., Benton, J., Bimaye, K., Blankenburg, K., Bonnin, D.,
Bouck, J., Bowie, S., Brieva, M., Brown, E., Brown, M., Bryant, N.P.,
Buha, C., Burch, P., Burkett, C., Burrell, K.L., Byrd, N.C.,
Carton, T.F., Carter, M., Cavazos, S.R., Chacko, J., Chavez, D.,
Chen, G., Chen, R., Chen, Z., Chowdhry, I., Christopoulos, C.,
Cleveland, C.D., Cox, C., Coyle, M.D., Dathorne, S.R., David, R.,
Devilla, M.L., Davis, C., Davy-Carroll, L., Dederich, D.A.,
Delaney, K.R., Delgado, O., Denn, A.L., Ding, Y., Dinh, H.H.,
Douthwaite, K.J., Draper, H., Dugan-Rocha, S., Durbin, K.J.,
Earnhart, C., Edgar, D., Edwards, C.C., Elhaj, C., Escotto, M.,
Falls, T., Ferraguto, D., Flagg, N., Ford, J., Foster, P., Frantz, P.,
Gabisi, A., Gao, J., Garcia, A., Garner, T., Garza, N., Gill, R.,
Correll, J.H., Guevara, W., Gunaratne, P., Hale, S., Hamilton, K.,
Harris, C., Harris, K., Hart, M., Havlak, P., Hawes, A., Hernandez, J.,
Hernandez, O., Hodgson, A., Hogues, M., Holloway, C., Hollins, B.,
Homs, F., Howard, S., Huber, J., Hulyk, S., Hume, J., Jackson, L.E.,
Jacobson, B., Jia, Y., Johnson, R., Jolivet, S., Joudah, S.,
Karlisson, E., Kelly, S., Khan, U., King, L., Korvah, J., Kovar, C.,

```

```

Kratovic, J., Kureshi, A., Landry, N., Leal, B., Lewis, L.C., Lewis, L.,
Li, J., Li, Z., Lichtarge, O., Lieu, J., Liu, J., Liu, W., Louissegh, H.,
Lozano, R.J., Lu, X., Lucier, A., Lucia, R., Luna, R., Ma, J.,
Maheshwari, M., Mapua, P., Martin, R., Martindale, A., Martinez, E.,
Massey, E., Mawhiney, E., McLeod, M.P., Meador, M., Mei, G., Metzker, M.,
Miner, G., Miner, Z., Mitchell, T., Mohabbat, K., Morgan, M., Morris, S.,
Moser, M., Neal, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N.,
Nguyen, N., Nickerson, E., Nwokenkwo, S., Ogih, M., Okwuonu, G.,
Oragunye, N., Oviedo, R., Pace, A., Payton, B., Peery, J., Perez, L.,
Peters, L., Pickens, R., Primus, E., Pu, L.L., Quiles, M., Ren, Y.,
Rives, M., Rojas, A., Rojubokan, I., Rolfe, M., Ruiz, S., Savary, G.,
Scherer, S., Scott, G., Shen, H., Shoohtari, N., Sisson, I.,
Sodergren, E., Sonaika, T., Sparks, A., Stanley, H., Stone, H.,
Sutton, A., Svatek, A., Tabor, P., Tamerisa, A., Tamerisa, K., Tang, H.,
Tansey, J., Taylor, C., Taylor, T., Telford, B., Thomas, N., Thomas, S.,
Usmani, K., Vazquez, L., Vera, V., Villalon, D., Vinson, R., Wang, Q.,
Wang, S., Ward-Moore, S., Warren, R., Washington, C., Watlington, S.,
Williams, G., Williamson, A., Wleciyk, R., Woodson, S., Worley, K.,
Wu, C., Wu, Y., Wu, Y.F., Zhou, J., Zorrilla, S., Nelson, D.,
Weinstock, G., and Gibbs, R.
Direct Submission
Unpublished
2 (bases 1 to 154265)
Worley, K.C.
Direct Submission
Submitted (19-JUL-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: KAAM
Center clone name: CH230-236P17
----- Summary Statistics
Sequencing vector: Plasmid;
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 115969 bases at least Q40
Consensus quality: 121810 bases at least Q30
Consensus quality: 127454 bases at least Q20
-----
* NOTE: Estimated insert size may differ from sequence length
(see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 61 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1
* 1224: contig of 1224 bp in length
* 1225
* 1324: gap of unknown length
* 1325
* 2862: contig of 1538 bp in length
* 2863
* 2962: gap of unknown length
* 2963
* 4039: contig of 1077 bp in length
* 4040
* 4139: gap of unknown length
* 4140
* 5220: contig of 1081 bp in length
* 5221
* 5320: gap of unknown length
* 5321
* 6757: contig of 1436 bp in length
* 6758
* 6857: gap of unknown length
* 6858
* 8352: contig of 1496 bp in length
* 8353
* 10449: contig of unknown length
* 10450
* 10549: contig of 1997 bp in length
* 10550
* 12609: gap of unknown length
* 12610
* 12709: contig of 2060 bp in length
* 12710
* 14051: gap of unknown length
* 14052
* 14151: contig of 1342 bp in length
* 14152
* 15296: contig of 1145 bp in length

```


Karlsson, E., Kelly, S., Khan, U., King, L., Korvah, J., Kovar, C.,
 Kratochvil, J., Kureshi, A., Landry, N., Leal, B., Lewis, L.C., Lewis, L.,
 Li, J., Li, Z., Lichtarge, O., Lieu, C., Liu, J., Liu, W., Louised, H.,
 Lozano, R.J., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J.,
 Maheshwari, M., Mapua, P., Martin, R., Martindale, A., Martinez, E.,
 Massey, E., Mawhney, E., McLeod, M.P., Meador, M., Mei, G., Metzger, M.,
 Miner, G., Miner, Z., Mitchell, T., Mohabbat, K., Morgan, M., Morris, S.,
 Moser, M., Neal, D., Newton, J., Newton, N., Nguyen, A., Nguyen, S.,
 Nguyen, N., Nickerson, E., Nwokenwo, S., Ogih, M., Okwuonu, G.,
 Oreguene, N., Ovioldo, R., Pace, A., Payton, B., Peery, J., Perez, L.,
 Peters, L., Pickens, R., Primus, E., Pu, L.L., Quiles, M., Ren, Y.,
 Rives, M., Rojas, A., Rojibokan, I., Rolfe, M., Ruiz, S., Savery, G.,
 Scherer, S., Scott, G., Shen, H., Shooshitari, N., Sisson, I.,
 Sodergren, E., Soraker, F., Sparks, A., Stanley, H., Stone, H.,
 Sutton, A., Svatek, A., Tabor, P., Tamerisa, A., Tamerisa, K., Tang, H.,
 Tansey, J., Taylor, C., Taylor, T., Telford, B., Thomas, N., Thomas, S.,
 Usman, K., Vasquez, L., Vera, V., Villalon, D., Vinson, R., Wang, Q.,
 Wang, S., Ward-Moore, S., Warren, R., Washington, C., Watlington, S.,
 Williams, G., Williamson, A., Wlezyk, R., Wooden, S., Worley, K.,
 Wu, C., Wu, Y., Wu, Y.F., Zhou, J., Zorrilla, S., Nelson, D.,
 Weinstein, G., and Gibbs, R.

JOURNAL

REFERENCE
 2 (bases 1 to 158286)
 Worley, K.C.

JOURNAL

Submitted (19-FEB-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 3 (bases 1 to 158286)
 Worley, K.C.

JOURNAL

Submitted (13-JUL-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 On Jul 12, 2002 this sequence version replaced gi:18701102.

COMMENT

Center: Baylor College of Medicine
 Center code: BCM
 Web site: <http://www.hgsc.bcm.tmc.edu/>
 Contact: hgsc-help@bcm.tmc.edu
 ----- Project Information
 Center project name: GMBI
 Center clone name: CH230-240A17
 ----- Summary Statistics
 Sequencing vector: Plasmid;
 Chemistry: Dye-terminator Big Dye; 100% of reads
 Assembly program: Phrap; version 0.990329
 Consensus quality: 103746 bases at least Q40
 Consensus quality: 108208 bases at least Q30
 Consensus quality: 111376 bases at least Q20

* NOTE: Estimated insert size may differ from sequence length
 (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
 * NOTE: This is a 'working draft' sequence. It currently
 consists of 68 contigs. The true order of the pieces
 is not known and their order in this sequence record is
 arbitrary. Gaps between the contigs are represented as
 runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence.
 * as soon as it is available and the accession number will
 be preserved.

1 1244: contig of 1244 bp in length
 1245 1344: gap of unknown length
 1345 2875: contig of 1531 bp in length
 2876 2975: gap of unknown length
 2976 4032: contig of 1057 bp in length
 4033 4132: gap of unknown length
 4133 5396: contig of 1264 bp in length
 5397 5496: gap of unknown length
 5497 6970: contig of 1473 bp in length
 6970 7070: gap of unknown length
 7070 8370: contig of 1301 bp in length

8371 8470: gap of unknown length
 8471 9974: contig of 1504 bp in length
 9975 10074: gap of unknown length
 10075 11969: contig of 1895 bp in length
 11970 12069: gap of unknown length
 12070 13421: contig of 1352 bp in length
 13422 13521: gap of unknown length
 13522 14649: contig of 1128 bp in length
 14650 14749: gap of unknown length
 14750 16351: contig of 1602 bp in length
 16352 16451: gap of unknown length
 16452 17940: contig of 1489 bp in length
 17941 18040: gap of unknown length
 18041 19287: contig of 1247 bp in length
 19288 19387: gap of unknown length
 19388 20768: contig of 1381 bp in length
 20769 20868: gap of unknown length
 20869 22395: contig of 1527 bp in length
 22396 22495: gap of unknown length
 22496 23637: contig of 1141 bp in length
 23637 25287: contig of 1551 bp in length
 25288 25387: gap of unknown length
 25388 26570: contig of 1183 bp in length
 26571 26670: gap of unknown length
 26671 28078: contig of 1408 bp in length
 28079 28178: gap of unknown length
 28179 29573: contig of 1395 bp in length
 29574 29673: gap of unknown length
 29674 31324: contig of 1651 bp in length
 31325 31424: gap of unknown length
 31425 32599: contig of 1175 bp in length
 32600 32699: gap of unknown length
 32700 34751: contig of 2052 bp in length
 34752 34851: gap of unknown length
 34852 36020: contig of 1169 bp in length
 36021 36120: gap of unknown length
 36121 37941: contig of 1821 bp in length
 37942 38041: gap of unknown length
 38042 39144: contig of 1103 bp in length
 39145 39244: gap of unknown length
 39245 41216: contig of 1971 bp in length
 41217 41315: gap of unknown length
 41316 43354: contig of 2039 bp in length
 43355 43454: gap of unknown length
 43455 45504: contig of 2050 bp in length
 45505 45604: gap of unknown length
 45605 47070: contig of 1466 bp in length
 47071 48306: contig of 1136 bp in length
 48307 48406: gap of unknown length
 48407 49467: contig of 1061 bp in length
 49468 49567: gap of unknown length
 49568 51036: contig of 1469 bp in length
 51037 51136: gap of unknown length
 51137 53712: contig of 2576 bp in length
 53713 53812: gap of unknown length
 53813 55162: contig of 1350 bp in length
 55163 55262: gap of unknown length
 55263 57095: contig of 1834 bp in length
 57096 57195: gap of unknown length
 57196 58767: contig of 1571 bp in length
 58768 58867: gap of unknown length
 58868 60411: contig of 1544 bp in length
 60412 60511: gap of unknown length
 60512 62118: contig of 1607 bp in length
 62119 62218: gap of unknown length
 62219 63943: contig of 1725 bp in length
 63944 64043: gap of unknown length
 64044 66201: contig of 2158 bp in length
 66202 68301: gap of unknown length
 68302 68867: contig of 2566 bp in length
 68868 68967: gap of unknown length

```

* 68968 71412: contig of 2445 bp in length
* 71413 71512: gap of unknown length
* 71513 74169: contig of 2657 bp in length
* 74170 74269: gap of unknown length
* 74270 78013: contig of 3744 bp in length
* 78014 78113: gap of unknown length
* 78114 80476: contig of 2363 bp in length
* 80477 80576: gap of unknown length
* 80577 83565: contig of 2989 bp in length
* 83566 83666: gap of unknown length
* 83667 85949: contig of 2284 bp in length
* 85950 86049: gap of unknown length
* 86050 89230: contig of 3201 bp in length
* 89231 91850: gap of unknown length
* 91851 91950: contig of 2500 bp in length
* 91951 91951: gap of unknown length
* 91952 93870: contig of 1920 bp in length
* 93871 93970: gap of unknown length
* 93971 96867: contig of 2897 bp in length
* 96868 96867: gap of unknown length
* 96968 99706: contig of 2739 bp in length
* 99707 99806: gap of unknown length
* 99807 101862: contig of 2056 bp in length

Query Match 58.5%; Score 90.4; DB 2; Length 158286;
Best Local Similarity 80.3%; Pred. No. 6.6e-15;
Matches 106; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

QY 1 ATGGCACACAACTAGACCTGGAAGAAATGCCAGCTGGATAGGCCAAGCCAGTGAAGCC 60
DB 76104 ATGGCAGACAGCAGCAGCATCGGAGAAATGCCAGCTTCGATAGGTCAGTCAAGTGAAGAAA 76163

QY 61 ACAGACATGCGAAGAACACTCTGATGACCAAGACGACGACGAGAGAGTGGAGT 120
DB 76164 ACTGACACATAGACAGACACCTGTGATGACCAAGAGACCATTTGAACAGAGAAAGAGGAGT 76223

QY 121 GAAATTTCTCGA 132
DB 76224 GAAATCTCTAA 76235

RESULT 7
AC129614/c
LOCUS
DEFINITION
Rattus norvegicus clone CH230-69C5, *** SEQUENCING IN PROGRESS ***,
75 unordered pieces.
AC129614
AC129614.2 GI:22091212
VERSION
KEYWORDS
SOURCE
ORGANISM
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.

REFERENCE
1 (bases 1 to 194568)
Muzny, D., Marle, Metzker, M., Lee, A., Abramson, S., Adams, C., Alder, J.,
Allen, C., Allen, H., Alsbrooks, S., Amin, A., Anguiano, D.,
Anyalebechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H.,
Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F.,
Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M.,
Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E.,
Cardenas, V., Carter, K., Cavazos, I., Cesar, H., Chen, X., Chen, Z., Chu, J.,
Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Y., D'Souza, L.,
Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L.,
Davila, M., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D.,
Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K.,
Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K.,
Egan, A., Escotto, M., Eugene, C., Evans, C.A., Falls, T., Fan, G.,
Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P.,
Fraser, C.M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M.,
Gebregorgis, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W.,
Gunaratne, P., Haaland, W., Hamil, C., Hamilton, C., Hamilton, K.,
Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J.,

```

```

Hernandez, R., Hines, S., Hladun, S.L., Hodgson, A., Hogues, M.,
Hollins, B., Howells, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A.,
Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A.,
Karpathy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C.,
Kowis, C., Kraft, C.L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J.,
Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J.,
Lorenshehwa, L., Louise, H., Lozada, R.J., Lu, X., Ma, J.,
Maheshwari, M., Mahindartne, M., Mahmoud, M., Malloy, K., Mangum, A.,
Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E.,
Mawhiney, S., McLeod, M., McNeill, T., Meenen, E., Milosavljevic, A.,
Miner, G., Minja, E., Montanayor, J., Moore, S., Morgan, M., Morris, K.,
Morris, S., Munidasa, M., Murphy, M., Nair, L., Nankervis, C., Neal, D.,
Newton, N., Nguyen, N., Norris, S., Nwaokemele, O., Okwuonu, G.,
Olarunpunsagoon, A., Pal, S., Parks, K., Pasternak, S., Paul, H.,
Perez, A., Perez, L., Pfannkuch, C., Plopper, F., Poindexter, A.,
Popovic, D., Primus, E., Pu, L.-L., Puazo, M., Quiroz, J., Rachlin, E.,
Reeves, K., Regier, M.A., Reigh, R., Reilly, B., Reilly, M., Ren, Y.,
Reuter, M., Richards, S., Riggs, F., Rives, C., Rodkey, T., Rojas, A.,
Rose, M., Rose, R., Ruiz, S.J., Sanders, W., Savary, G., Scherer, S.,
Scott, G., Shatsman, S., Shen, H., Shetty, J., Shvartsbeyn, A.,
Sisson, I., Sitter, C.D., Smajd, D., Sneed, A., Sodergren, E.,
Song, X.-Z., Sorelle, R., Sosa, J., Steimle, M., Strong, R., Sutton, A.,
Svatek, A., Tabor, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S.,
Tingey, A., Trejos, Z., Usmani, K., Valas, R., Vera, V., Villasana, D.,
Waldron, D., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J.,
Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wleczyk, R.,
Wooden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S.,
Yan, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X.,
Zhao, S., Dunn, D., von Niederhauser, A., Weiss, R., Smith, D.R.,
Holt, R.A., Smith, H.O., Weinstein, G. and Gibbs, R.A.

Direct Submission
Unpublished
2 (bases 1 to 194568)
Worley, K.C.
Direct Submission
Submitted (31-JUL-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 194568)
Rat Genome Sequencing Consortium.
Direct Submission
Submitted (24-AUG-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Aug 2, 2002 this sequence version replaced gi:22024297.

-----
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
-----
Project Information
Center project name: KBYZ
Center clone name: CH230-69C5
-----
Summary Statistics
Sequencing vector: Plasmid;
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 132737 bases at least Q40
Consensus quality: 140036 bases at least Q30
Consensus quality: 146272 bases at least Q20
-----
* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 75 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 1169: contig of 1169 bp in length
* 1170 1269: gap of unknown length

```



```

Contact: submissions@watson.wustl.edu
----- Project Information -----
Center project name: M.BB0212G16
----- Summary Statistics -----
Sequencing vector: M13; 0%
Chemistry: Dye-terminator; 100%
Assembly: Dye-terminator Big Dye; 100% of reads
Consensus quality: Phrap; version 0.990319
Consensus quality: 182274 bases at least Q40
Consensus quality: 182633 bases at least Q30
Consensus quality: 182875 bases at least Q20
* NOTE: This is a 'working draft' sequence. It currently
* consists of 3 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1 15774: contig of 15774 bp in length
* 15775 15874: gap of unknown length
* 15875 90246: contig of 74372 bp in length
* 90247 90346: gap of unknown length
* 90347 183583: contig of 93237 bp in length.
FEATURES
Location/Qualifiers
1. 183583
/organism="Mus musculus"
/db_xref="taxon:10090"
/chromosome="UNK"
/clone="RP24-212G16"
1. 15774
/note="assembly_name:Contig13"
15875..90246
/note="assembly_name:Contig14"
90347..183583
/note="assembly_name:Contig15"
BASE COUNT 47220 a 44434 c 45182 g 46547 t 200 others
ORIGIN

Query Match 67.3%; Score 88.8; DB 2; Length 183583;
Best Local Similarity 79.5%; Pred. No. 1.9e-14;
Matches 105; Conservative 0; Mismatches 27; Indels 0; Gaps 0;

QY 1 ATGGCACACAACTGACCTGGAGAAATTCGACGCTGGATAGGCGCAGCTCAAGGCC 60
Db 68234 ATGGCAGACACCGGACATAGGGGAAATTCGACGCTCAATAGGCGCAGCTGAAGAAA 68293

QY 61 ACAGAGATGCGAGAGAACACTCTGATGACCAAGAGACACAGAGAGAGAGTGGAGT 120
Db 68294 ACCGAGCGCAGAGAGACACCTCGCGACCAAGAGACCACTGACAGGAAAGAGAGT 68353

QY 121 GAAATTTCTGTA 132
Db 68354 GAAATCTCTTAA 68365

RESULT 9
AC131561
LOCUS
DEFINITION Rattus norvegicus clone CH230-8p16, *** SEQUENCING IN PROGRESS ***,
72 unordered pieces.
ACCESSION AC131561
VERSION AC131561.1 GI:22474831
KEYWORDS HTG; HTGS_PHASE1.
SOURCE Norway rat.
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE 1 (bases 1 to 186744)

```

AUTHORS

Muzny, D. Marie., Metzker, M. Lee., Abramson, S., Adams, C., Alder, J., Allen, C., Allen, H., Alsbrooks, S., Amin, A., Anguiano, D., Anyalebechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H., Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F., Biswal, K., Blair, J., Burch, P., Burrell, K., Blyth, P., Brown, M., Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E., Cardenas, V., Carter, K., Cavazos, I., Cesar, H., Chen, A., Chu, J., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, J., Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L., Davila, M. L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D., Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K., Egan, A., Escotto, M., Eugene, C., Evans, C. A., Falls, T., Fan, G., Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P., Fraser, C. M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M., Gregorogis, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W., Gunaratne, P., Haaland, W., Hamill, C., Hamilton, C., Hamilton, K., Harvey, I., Havlak, P., Hawes, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladun, S. L., Hodgson, A., Hogue, M., Hollins, B., Howells, S., Rulyk, S., Hume, J., Idlebird, D., Jackson, A., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A., Karpathy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C., Kowis, C., Kraft, C. L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Lorensuhera, L., Loulseged, H., Lozano, R. J., Lu, X., Ma, J., Maheshwari, M., Mahindartne, M., Mahmoud, M., Malloy, K., Mangum, A., Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E., Mawhiney, S., McLeod, M., McNeill, T., Meenen, E., Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nwaokemele, O., Okwuonu, G., Olarunpunsagoon, A., Pal, S., Parks, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C., Plopper, F., Poindexter, A., Popovic, D., Primus, E., Pu, L., Puazo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M. A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S. J., Sanders, W., Savory, G., Scherer, S., Scott, G., Shatsman, S., Shen, H., Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C. D., Smajs, D., Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J., Steimle, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K., Valas, R., Vera, V., Villasana, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Willison, R., Wleczek, R., Wooten, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von Niederhausen, A., Weiss, R., Smith, D. R., Holt, R. A., Smith, H. O., Weinstein, G. and Gibbs, R. A.

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
----- Project Information -----
Center project name: GCXJ
Center clone name: CH230-8p16
----- Summary Statistics -----
Sequencing vector: Plasmid;
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 121943 bases at least Q40
Consensus quality: 127700 bases at least Q30
Consensus quality: 132440 bases at least Q20

* NOTE: Estimated insert size may differ from sequence length

AUTHORS	JOURNAL	REFERENCE	AUTHORS	JOURNAL	REFERENCE	AUTHORS	JOURNAL	COMMENT
---------	---------	-----------	---------	---------	-----------	---------	---------	---------

```

----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc/Index.shtml
Contact: submissions@wustl.edu
----- Project Information -----
Center project name: M_BA0389K05
----- Summary Statistics -----
Sequencing vector: M13; 0%
Sequencing vector: plasmid; 100%
Chemistry: dye-primer ET; 0% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 326977 bases at least Q40
Consensus quality: 329212 bases at least Q30
Consensus quality: 330686 bases at least Q20
Insert size: 172000; agarose-fp
Insert size: 335212; sum-of-contigs
Quality coverage: 13.26 in Q20 bases; agarose-fp
Quality coverage: 7.79 in Q20 bases; sum-of-contigs

```

<pre>/organism="Mus musculus"</pre>					
<pre>/db_xref="taxon:10090"</pre>					
<pre>/chromosome="UNK"</pre>					
<pre>/clone="RP23-389KS"</pre>					
<pre>1..2398</pre>					
<pre>note="assembly_name:Contig15"</pre>					
<pre>2499..5211</pre>					
<pre>note="assembly_name:Contig16"</pre>					
<pre>5312..9003</pre>					
<pre>note="assembly_name:Contig17"</pre>					
<pre>9104..16087</pre>					
<pre>note="assembly_name:Contig18"</pre>					
<pre>16188..22311</pre>					
<pre>note="assembly_name:Contig19"</pre>					
<pre>22412..31922</pre>					
<pre>note="assembly_name:Contig20"</pre>					
<pre>32023..43567</pre>					
<pre>note="assembly_name:Contig21"</pre>					
<pre>43668..59035</pre>					
<pre>note="assembly_name:Contig22"</pre>					
<pre>59136..75963</pre>					
<pre>note="assembly_name:Contig23"</pre>					
<pre>76064..101243</pre>					
<pre>note="assembly_name:Contig24"</pre>					
<pre>101344..129536</pre>					
<pre>note="assembly_name:Contig25"</pre>					
<pre>129637..152205</pre>					
<pre>note="assembly_name:Contig26"</pre>					
<pre>152306..222216</pre>					
<pre>note="assembly_name:Contig27"</pre>					
<pre>222317..336512</pre>					
<pre>note="assembly_name:Contig28"</pre>					
<pre>BASE COUNT 85482 a 84197 c 83135 g 82393 t 1305 others</pre>					
<pre>ORIGIN</pre>					
<pre>Query Match 67.3%; Score 88.8; DB 2; Length 336512;</pre>					
<pre>Best Local Similarity 79.5%; Pred. No. 1.9e-14;</pre>					
<pre>Matches 105; Conservative 0; Mismatches 27; Indels 0; Gaps 0;</pre>					
Qy	1	ATGCACACAACCTAGACTTGGAGAATAATTGCCAGCTTGATTAAGGCGCAAGCTGAAGGCC	60		
Db	204670	ATGCAGACAAACCCGACATAGGGGANAATFCGCCAGCTTCANRANGGCAAGCTGAAGAAA	204728		
Qy	61	ACAGAGATGCAAGAGAACAACCTCTGTATGACCAAGAGACCACAGCAGCAGGAGTAGTGAGT	120		
Db	204730	ACCGAGACGAGGAGAACACCTCGCGACCAAGAGACCACCTGAACAGGAAAAAGAGAGT	204785		
Qy	121	GAAATTTCTCTGA	132		
Db	204790	GAAATCTCTTAA	204801		
<pre>RESULT 11</pre>					
G06175	G06175	human STS WI-6018, sequence tagged site.	174 bp	DNA	linear STS 19-OCT-1995
LOCUS	DEFINITION				
ACCESSION	G06175				
VERSION	G06175.1	GI-859420			
KEYWORDS	STS; STS sequence; primer; sequence tagged site.				
SOURCE	Homo sapiens STS derived from sequences in dbEST and the Unigene collection.				
ORGANISM	Homo sapiens				
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
REFERENCE	1 (bases 1 to 174)				
AUTHORS	Hudson,T.				
TITLE	Whitehead Institute/MIT Center for Genome Research; Physically				
JOURNAL	Mapped ESTs				
COMMENT	Unpublished (1995)				
<pre>Contact: Thomas Hudson Whitehead Institute/MIT Center for Genome Research Whitehead Institute for Biomedical Research</pre>					

9 Cambridge Center, Cambridge MA 02142 USA
Tel: 617 252 1900
Fax: 617 252 1902
Email: thudson@genome.wi.mit.edu

Primer A: CACGATGCTTTTAAAGAAATGG
Primer B: GAAATTCACCTCGCTTCTCC
STS size: 150
PCR Profile:

Presoak:

Denaturation:

Annealing: 56 degrees C

Polymerization:

PCR Cycles: 35

Thermal Cycler:

Protocol:

Template: 10 ng

Primer: each 5 pm

dNTPs: each 4 mM

Taq Polymerase: 0.025 units/ul

Total Vol: 20 ul

Buffer:

MgCl2: 1.5 mM

KCl: 50 mM

Tris-HCL: 10 mM

pH: 9.3

Prepared with primer pairs derived from T27818 -- dbEST.
Location/Qualifiers
1. .174

/organism="Homo sapiens"

/db_xref="taxon:9606"

/map="808_E_1; 849_E_2; 931_E_8; (789-796)_C_5;

752_C_(2,7,11)"

20. .169

20. .142

Complement(149. .169)

62 a 37 c 52 g 23 t

FEATURES

source

LOCUS

AX381579

DEFINITION

Sequence 517 from Patent WO0212280.

ACCESSION

AX381579

VERSION

AX381579.1

KEYWORDS

GI:19576398

SOURCE

human.

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

Pyle, R.A., Xu, J., and Secrist, H.

Compositions and methods for the therapy and diagnosis of colon

cancer

Patent WO0212280.

Linear

PAT 18-MAR-2002

243 bp

DNA

Indels 3; Gaps 1;

Mismatches 19;

Length 174;

Score 88.6; DB 11;

Pred. No. 1.3e-14;

Conservative 0;

Best Local Similarity 83.7%;

Score 88.6; DB 11;

Pred. No. 1.3e-14;

Conservative 0;

Best Local Similarity 83.7%;

Score 88.6; DB 11;

Pred. No. 1.3e-14;

JOURNAL

source

LOCUS

AX381269

DEFINITION

Sequence 224 from Patent WO0212280.

ACCESSION

AX381269

VERSION

AX381269.1

KEYWORDS

GI:19576105

SOURCE

human.

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

Pyle, R.A., Xu, J., and Secrist, H.

Compositions and methods for the therapy and diagnosis of colon

cancer

Patent WO0212280.

Linear

PAT 18-MAR-2002

353 bp

DNA

Indels 3; Gaps 1;

Mismatches 19;

Length 243;

Score 88.6; DB 6;

Pred. No. 1.3e-14;

Conservative 0;

Best Local Similarity 83.7%;

Score 88.6; DB 6;

Pred. No. 1.3e-14;

Conservative 0;

Best Local Similarity 83.7%;

Score 88.6; DB 6;

Pred. No. 1.3e-14;

Conservative 0;

Best Local Similarity 83.7%;

Score 88.6; DB 6;

Pred. No. 1.3e-14;

Conservative 0;

Best Local Similarity 83.7%;

Score 88.6; DB 6;

Pred. No. 1.3e-14;

Conservative 0;

Best Local Similarity 83.7%;

Score 88.6; DB 6;

Pred. No. 1.3e-14;

Conservative 0;

Best Local Similarity 83.7%;

Score 88.6; DB 6;

Pred. No. 1.3e-14;

Conservative 0;

Best Local Similarity 83.7%;

Score 88.6; DB 6;

Pred. No. 1.3e-14;

KEYWORDS

SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE 1
Pyle,R.A., Xu,J. and Secrist,H.
JOURNAL Compositions and methods for the therapy and diagnosis of colon
PATENT: WO 0212280-A 207 14-FEB-2002;
CORIXA CORPORATION (US)
FEATURES Location/Qualifiers
source
1..395
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 105 a 113 c 115 g 62 t
ORIGIN

Query Match 67.1%; Score 88.6; DB 6; Length 395;
Best Local Similarity 83.7%; Pred. No. 1.4e-14;
Matches 113; Conservative 0; Mismatches 19; Indels 3; Gaps 1;
QY 1 ATGGCAGACAACTGAGCAATTCGCCAGCTTGGATAGGCCCAAGCC 60
DB 72 ATGGCAGACAACTGAGCAATTCGCCAGCTTGGATAGGCCCAAGCC 131
QY 61 ACAGAGATGC---AGAAGAACACTCTGATGACCAAGAGACCAAGAGAGAGAGTGG 117
DB 132 ACGGAGACGAGAGAGAGACACCTTCCGACCAAGAGAGACCTTGGAGGAGAGCGG 191
QY 118 AGTGAATTTCTCTGA 132
DB 192 AGTGAATTTCTCTAA 206

RESULT 15

HUMTHMBX
LOCUS Human thymosin beta 10 mRNA, complete cds. PRI 03-AUG-1993
DEFINITION
ACCESSION M92381
VERSION M92381.1 GI:339660
KEYWORDS thymosin beta 10.
SOURCE Homo sapiens cDNA to mRNA.
ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE 1 (bases 1 to 400)
HALL,A.K., Hempstead,J. and Morgan,J.I.
Thymosin beta 10 levels in developing human brain and its
JOURNAL regulation by retinoic acid in the HTB-10 neuroblastoma
MEDLINE Brain Res. Mol. Brain Res. 8 (2), 129-135 (1990)
PUBMED 90384336
PUBMED 2169566

FEATURES

Location/Qualifiers
source
1..400
/organism="Homo sapiens"
/db_xref="taxon:9606"
36..152
/codon_start=1
/product="thymosin beta 10"
/protein_id="AAC41691.1"
/db_xref="GI:339661"
/translation="MGEIASFDKALKTKTQEKNTLPTKETIEQKRSEIS"
BASE COUNT 118 a 106 c 105 g 71 t
ORIGIN

Query Match 67.1%; Score 88.6; DB 9; Length 400;
Best Local Similarity 83.7%; Pred. No. 1.4e-14;
Matches 113; Conservative 0; Mismatches 19; Indels 3; Gaps 1;
QY 1 ATGGCAGACAACTGAGCAATTCGCCAGCTTGGATAGGCCCAAGCC 60
DB 18 ATGGCAGACAACTGAGCAATTCGCCAGCTTGGATAGGCCCAAGCC 77

QY 61 ACAGAGATGC---AGAAGAACACTCTGATGACCAAGAGACCAAGAGAGAGTGG 117
DB 78 ACGGAGACGAGAGAGAGACACCTTCCGACCAAGAGAGACCTTGGAGGAGAGCGG 137
QY 118 AGTGAATTTCTCTGA 132
DB 138 AGTGAATTTCTCTAA 152
Search completed: June 2, 2003, 14:36:33
Job time : 1077 secs

THIS PAGE BLANK

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 30, 2003, 10:34:51 ; Search time 44 Seconds
(without alignments)
98.923 Million cell updates/sec

Title: US-09-915-178-2

Perfect score: 213

Sequence: 1 MAHKLDLEIASLDKAKLKA.....QKNTLMKTETQEKWSEIS 43

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 383519 seqs, 101223694 residues

Total number of hits satisfying chosen parameters: 383519

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications_AA:*

- 1: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
- 2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
- 10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
- 12: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*
- 13: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
- 14: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	135.5	63.6	49	9 US-09-842-758-51	Sequence 51, Appl
2	130.5	61.3	43	9 US-09-772-445A-7	Sequence 7, Appl
3	116.5	54.7	41	9 US-09-772-445A-11	Sequence 11, Appl
4	112.5	52.8	41	9 US-09-772-445A-5	Sequence 5, Appl
5	110.5	51.9	41	9 US-09-772-445A-6	Sequence 6, Appl
6	106.5	50.0	41	9 US-09-772-445A-8	Sequence 8, Appl
7	104.5	49.1	40	9 US-09-772-445A-15	Sequence 15, Appl
8	99.5	46.7	40	9 US-09-772-445A-12	Sequence 12, Appl
9	99.5	46.7	42	9 US-09-772-445A-9	Sequence 9, Appl
10	97.5	45.8	43	9 US-09-772-445A-4	Sequence 4, Appl
11	97.5	45.8	44	9 US-10-171-311-220	Sequence 220, Appl
12	97.5	45.8	50	9 US-09-842-758-50	Sequence 50, Appl
13	97.5	45.8	56	9 US-09-925-299-1252	Sequence 1252, Appl
14	97.5	45.8	56	9 US-09-842-758-49	Sequence 49, Appl
15	97.5	45.8	56	10 US-09-925-299-1252	Sequence 1252, Appl
16	96.5	45.3	43	9 US-09-772-445A-10	Sequence 10, Appl
17	95.5	44.8	45	9 US-09-874-736-2	Sequence 2, Appl
18	95.5	44.8	45	9 US-09-842-758-52	Sequence 52, Appl
19	93.5	43.9	40	9 US-09-772-445A-14	Sequence 14, Appl

20	91.5	43.0	45	9 US-09-842-758-48	Sequence 48, Appl
21	90.5	42.5	43	9 US-09-772-445A-3	Sequence 3, Appl
22	90.5	42.5	44	9 US-09-772-445A-13	Sequence 13, Appl
23	89.5	42.0	38	10 US-09-879-666-6	Sequence 6, Appl
24	87.5	41.1	43	9 US-09-772-445A-2	Sequence 2, Appl
25	69	32.4	69	9 US-09-842-758-10	Sequence 10, Appl
26	64	30.0	120	9 US-09-746-783-190	Sequence 190, Appl
27	61	28.6	517	10 US-09-815-242-4970	Sequence 4970, Appl
28	61	28.6	518	10 US-09-815-242-10891	Sequence 10891, Appl
29	54	25.4	289	10 US-09-764-864-911	Sequence 911, Appl
30	54	25.4	388	10 US-09-764-864-931	Sequence 931, Appl
31	54	25.4	739	9 US-09-874-162A-8	Sequence 5, Appl
32	54	25.4	776	9 US-09-874-162A-8	Sequence 8, Appl
33	52	24.4	122	9 US-09-897-645-1	Sequence 1, Appl
34	52	24.4	122	10 US-09-815-242-13472	Sequence 13472, Appl
35	52	24.4	122	10 US-09-815-242-13649	Sequence 13649, Appl
36	52	24.4	418	9 US-09-893-519A-10	Sequence 10, Appl
37	52	24.4	1530	9 US-10-118-513A-6	Sequence 6, Appl
38	52	24.4	2266	9 US-10-118-513A-14	Sequence 14, Appl
39	51.5	24.2	174	10 US-09-864-761-48360	Sequence 48360, Appl
40	51.5	24.2	533	10 US-09-815-242-11612	Sequence 11612, Appl
41	51	23.9	488	9 US-10-033-297-141	Sequence 141, Appl
42	51	23.9	488	9 US-09-940-244-141	Sequence 141, Appl
43	51	23.9	888	9 US-10-231-035-3	Sequence 3, Appl
44	51	23.9	1207	10 US-09-927-112-2	Sequence 2, Appl
45	50.5	23.7	458	9 US-09-965-529-2	Sequence 2, Appl

ALIGNMENTS

RESULT 1

US-09-842-758-51

Sequence 51, Application US/09842758

Publication No. US20030083244A1

GENERAL INFORMATION:

APPLICANT: Vernet, Corine A. M.
 APPLICANT: Fernandes, Elma R.
 APPLICANT: Gerlach, Valerie
 APPLICANT: Shimkets, Richard A.
 APPLICANT: Malyakkar, Uriel M.
 APPLICANT: Boldog, Ferenc L.
 APPLICANT: Zerkhusen, Bryan D.
 APPLICANT: Spytek, Kimberly A.
 APPLICANT: Majumder, Kumud
 APPLICANT: Tchernev, Velizar T.
 APPLICANT: Padigaru, Muralidhara
 APPLICANT: Patturajan, Meera
 APPLICANT: Burgess, Catherine E.
 APPLICANT: Gangolli, Esha A.
 APPLICANT: Smithson, Glennnda
 APPLICANT: Rastelli, Luca
 APPLICANT: MacDougall, John R.
 APPLICANT: Taupier, Raymond J.
 APPLICANT: Grosse, William M.
 APPLICANT: Edward, Szekeres S.
 APPLICANT: Alsobrook II, John P.
 TITLE OF INVENTION: No. US20030083244A1el Proteins and Nucleic Acids Encoding Sa
 FILE REFERENCE: 15966-783
 CURRENT APPLICATION NUMBER: US/09/842,758
 CURRENT FILING DATE: 2001-04-25
 PRIOR APPLICATION NUMBER: 60/200,158
 PRIOR FILING DATE: 2000-04-26
 PRIOR APPLICATION NUMBER: 60/200,613
 PRIOR FILING DATE: 2000-04-28
 PRIOR APPLICATION NUMBER: 60/201,006
 PRIOR FILING DATE: 2000-05-01
 PRIOR APPLICATION NUMBER: 60/201,007
 PRIOR FILING DATE: 2000-05-01
 PRIOR APPLICATION NUMBER: 60/201,236
 PRIOR FILING DATE: 2000-05-01

```

; PRIOR APPLICATION NUMBER: 60/201,238
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: 60/201,186
; PRIOR FILING DATE: 2000-05-02
; PRIOR APPLICATION NUMBER: 60/201,474
; PRIOR FILING DATE: 2000-05-03
; PRIOR APPLICATION NUMBER: 60/201,508
; PRIOR FILING DATE: 2000-05-03
; PRIOR APPLICATION NUMBER: 60/220,591
; PRIOR FILING DATE: 2000-07-25
; PRIOR APPLICATION NUMBER: 60/232,678
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: 60/263,217
; PRIOR FILING DATE: 2001-01-22
; PRIOR APPLICATION NUMBER: 60/265,160
; PRIOR FILING DATE: 2001-01-30
; NUMBER OF SEQ ID NOS: 113
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 51
; LENGTH: 49
; TYPE: PRT
; ORGANISM: Homo sapiens
; ORGANISM: Homo sapiens
US-09-842-758-51

Query Match 63.6%; Score 135.5; DB 9; Length 49;
Best Local Similarity 75.0%; Pred. No. 4e-10; Indels 1; Gaps 1;
Matches 33; Conservative 1; Mismatches 9; Indels 1; Gaps 1;

QY 1 MAHKLDLEIASLDRKAKLKATEMQ-KNTLMTKETTEQEKWSEIS 43
||| : ||||| ||| ||| ||||| ||||| |||||
DB 6 MADKPDGMEIASFDRAKLKKTETQEKNTLPTKETTEQEKRSSEIS 49

RESULT 2
US-09-772-445A-7
; Sequence 7, Application US/09772445A
; Publication No. US20030060405A1
; GENERAL INFORMATION:
; APPLICANT: Kleinman, Hynda K.
; APPLICANT: Goldstein, Allan L.
; APPLICANT: Malinda, Katherine M.
; APPLICANT: Sosne, Gabriel
; TITLE OF INVENTION: THYMOSIN BETA 4 PROMOTES WOUND REPAIR
; FILE REFERENCE: 08830-056001
; CURRENT APPLICATION NUMBER: US/09/772,445A
; CURRENT FILING DATE: 2001-01-29
; PRIOR APPLICATION NUMBER: PCT/US99/17282
; PRIOR FILING DATE: 1999-07-30
; PRIOR APPLICATION NUMBER: US 60/094,690
; PRIOR FILING DATE: 1998-07-30
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 43
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-772-445A-7

Query Match 61.3%; Score 130.5; DB 9; Length 43;
Best Local Similarity 74.4%; Pred. No. 1.4e-09; Indels 1; Gaps 1;
Matches 32; Conservative 1; Mismatches 9; Indels 1; Gaps 1;

QY 2 AHKLDLEIASLDRKAKLKATEMQ-KNTLMTKETTEQEKWSEIS 43
||| : ||||| ||| ||| ||||| ||||| |||||
DB 1 ADKPDGMEIASFDRAKLKKTETQEKNTLPTKETTEQEKRSSEIS 43

RESULT 3
US-09-772-445A-11
; Sequence 11, Application US/09772445A
; Publication No. US20030060405A1
; GENERAL INFORMATION:
; APPLICANT: Kleinman, Hynda K.
; APPLICANT: Goldstein, Allan L.
; APPLICANT: Malinda, Katherine M.
; APPLICANT: Sosne, Gabriel
; TITLE OF INVENTION: THYMOSIN BETA 4 PROMOTES WOUND REPAIR
; FILE REFERENCE: 08830-056001
; CURRENT APPLICATION NUMBER: US/09/772,445A
; CURRENT FILING DATE: 2001-01-29
; PRIOR APPLICATION NUMBER: PCT/US99/17282
; PRIOR FILING DATE: 1999-07-30
; PRIOR APPLICATION NUMBER: US 60/094,690
; PRIOR FILING DATE: 1998-07-30
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 41
; TYPE: PRT
; ORGANISM: Bos taurus
US-09-772-445A-5

Query Match 52.8%; Score 112.5; DB 9; Length 41;
Best Local Similarity 73.7%; Pred. No. 2.3e-07; Indels 1; Gaps 1;
Matches 28; Conservative 0; Mismatches 9; Indels 1; Gaps 1;

QY 2 AHKLDLEIASLDRKAKLKATEMQ-KNTLMTKETTEQEK 38
||| : ||||| ||| ||| ||||| ||||| |||||
DB 1 ADKPDGMEIASFDRAKLKKTETQEKNTLPTKETTEQEK 38

RESULT 5
US-09-772-445A-6
; Sequence 6, Application US/09772445A
; Publication No. US20030060405A1
; GENERAL INFORMATION:
; APPLICANT: Kleinman, Hynda K.
; APPLICANT: Goldstein, Allan L.
; APPLICANT: Malinda, Katherine M.
; APPLICANT: Sosne, Gabriel
; TITLE OF INVENTION: THYMOSIN BETA 4 PROMOTES WOUND REPAIR
; FILE REFERENCE: 08830-056001
; CURRENT APPLICATION NUMBER: US/09/772,445A
; CURRENT FILING DATE: 2001-01-29
; PRIOR APPLICATION NUMBER: PCT/US99/17282
; PRIOR FILING DATE: 1999-07-30
; PRIOR APPLICATION NUMBER: US 60/094,690
; PRIOR FILING DATE: 1998-07-30
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 41
; TYPE: PRT
; ORGANISM: Bos taurus
US-09-772-445A-5

Query Match 54.7%; Score 116.5; DB 9; Length 41;
Best Local Similarity 73.7%; Pred. No. 7.5e-08; Indels 1; Gaps 1;
Matches 28; Conservative 1; Mismatches 8; Indels 1; Gaps 1;

QY 2 AHKLDLEIASLDRKAKLKATEMQ-KNTLMTKETTEQEK 38
||| : ||||| ||| ||| ||||| ||||| |||||
DB 1 ADKPDGMEIASFDRAKLKKTETQEKNTLPTKETTEQEK 38

RESULT 4
US-09-772-445A-5
; Sequence 5, Application US/09772445A
; Publication No. US20030060405A1
; GENERAL INFORMATION:
; APPLICANT: Kleinman, Hynda K.
; APPLICANT: Goldstein, Allan L.
; APPLICANT: Malinda, Katherine M.
; APPLICANT: Sosne, Gabriel
; TITLE OF INVENTION: THYMOSIN BETA 4 PROMOTES WOUND REPAIR
; FILE REFERENCE: 08830-056001
; CURRENT APPLICATION NUMBER: US/09/772,445A
; CURRENT FILING DATE: 2001-01-29
; PRIOR APPLICATION NUMBER: PCT/US99/17282
; PRIOR FILING DATE: 1999-07-30
; PRIOR APPLICATION NUMBER: US 60/094,690
; PRIOR FILING DATE: 1998-07-30
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 41
; TYPE: PRT
; ORGANISM: Bos taurus
US-09-772-445A-5

Query Match 52.8%; Score 112.5; DB 9; Length 41;
Best Local Similarity 73.7%; Pred. No. 2.3e-07; Indels 1; Gaps 1;
Matches 28; Conservative 0; Mismatches 9; Indels 1; Gaps 1;

QY 2 AHKLDLEIASLDRKAKLKATEMQ-KNTLMTKETTEQEK 38
||| : ||||| ||| ||| ||||| ||||| |||||
DB 1 ADKPDGMEIASFDRAKLKKTETQEKNTLPTKETTEQEK 38

RESULT 5
US-09-772-445A-6
; Sequence 6, Application US/09772445A
; Publication No. US20030060405A1
; GENERAL INFORMATION:
; APPLICANT: Kleinman, Hynda K.
; APPLICANT: Goldstein, Allan L.
; APPLICANT: Malinda, Katherine M.
; APPLICANT: Sosne, Gabriel
; TITLE OF INVENTION: THYMOSIN BETA 4 PROMOTES WOUND REPAIR
; FILE REFERENCE: 08830-056001
; CURRENT APPLICATION NUMBER: US/09/772,445A
; CURRENT FILING DATE: 2001-01-29
; PRIOR APPLICATION NUMBER: PCT/US99/17282
; PRIOR FILING DATE: 1999-07-30
; PRIOR APPLICATION NUMBER: US 60/094,690
; PRIOR FILING DATE: 1998-07-30
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 41
; TYPE: PRT
; ORGANISM: Balaenoptera acutorostrata
US-09-772-445A-11

Query Match 54.7%; Score 116.5; DB 9; Length 41;
Best Local Similarity 73.7%; Pred. No. 7.5e-08; Indels 1; Gaps 1;
Matches 28; Conservative 1; Mismatches 8; Indels 1; Gaps 1;

QY 2 AHKLDLEIASLDRKAKLKATEMQ-KNTLMTKETTEQEK 38
||| : ||||| ||| ||| ||||| ||||| |||||
DB 1 ADKPDGMEIASFDRAKLKKTETQEKNTLPTKETTEQEK 38
```

1
 2
 3
 4
 5
 6
 7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525

: SEQ ID NO: 3
: LENGTH: 42

APPLICANT:	Fernandes, Elma R
APPLICANT:	Gerlach, Valerie
APPLICANT:	Shimkes, Richard A
APPLICANT:	Malyankar, Uriel M
APPLICANT:	Boldog, Ferenc L
APPLICANT:	Zerhusen, Bryan D
APPLICANT:	Sytek, Kimberly A
APPLICANT:	Muamder, Kumud
APPLICANT:	Tchernev, Velizar T
APPLICANT:	Padigaru, Muralidhara
APPLICANT:	Patturajan, Meera
APPLICANT:	Burgess, Catherine E
APPLICANT:	Gangolli, Esha A
APPLICANT:	Smithson, Glenda
APPLICANT:	Rastelli, Luca
APPLICANT:	MacDougall, John R
APPLICANT:	Taupier, Raymond J
APPLICANT:	Grosse, William M

```

RESULT 15
US-09-925-299-1252
; Sequence 1252, Application US/09925299
; Patent No. US20020055627A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA102
; CURRENT APPLICATION NUMBER: US/09/925,299
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05883
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1556
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1252
; LENGTH: 56
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-925-299-1252

```

us-09-915-178-2.rapb

	Query Match	45.88;	Score: 97.5;	DB 10;	Length 56;
	Best Local Similarity	59.08;	Pred. No. 2.5e-05;		
	Matches 23;	Conservative	4;	Mismatches 11;	Indels 1;
					Gaps 1;
QY	1	MAHKLDLEETASD	KAKLKAT	EMQ-KNTLMTKET	EQEK 38
db	13	MSDKPDMAELKEFD	KSLKAKT	FQEKRNPLP	KSETIEQEK 51

Search completed: May 30, 2003, 10:43:49
Job time : 45 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 30, 2003, 10:31:35 ; Search time 26 Seconds
(without alignments)
48.661 Million cell updates/sec

Title: US-09-915-178-2
Perfect score: 213
Sequence: 1 MAHKLDLEIASLDRKAKLKA.....QKTLMTKETTEQKRWSEIS 43

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2.6/prodata/1/1aa/5A_COMB.pep.*
2: /cgn2.6/prodata/1/1aa/5B_COMB.pep.*
3: /cgn2.6/prodata/1/1aa/6A_COMB.pep.*
4: /cgn2.6/prodata/1/1aa/6B_COMB.pep.*
5: /cgn2.6/prodata/1/1aa/PCTUS_COMB.pep.*
6: /cgn2.6/prodata/1/1aa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	97.5	45.8	44	3	US-09-058-489-27
2	95.5	44.8	45	1	US-08-664-856A-2
3	95.5	44.8	45	1	US-08-801-796-2
4	95.5	44.8	45	2	US-08-931-877-2
5	95.5	44.8	45	2	US-08-664-857A-2
6	95.5	44.8	45	3	US-09-069-484-2
7	95.5	44.8	45	4	US-09-135-599-2
8	95.5	44.8	45	4	US-09-369-744-2
9	88.5	41.5	44	3	US-09-058-489-28
10	54.5	25.6	245	1	US-08-317-707-2
11	54.5	25.6	245	2	US-08-514-921-2
12	54.5	25.6	245	2	US-08-514-921-4
13	54.5	25.6	245	2	US-08-514-921-6
14	54.5	25.6	245	2	US-08-514-921-8
15	53.5	25.1	210	4	US-08-858-207A-269
16	52.5	24.6	272	4	US-09-029-213B-27
17	52	24.4	122	3	US-08-987-144-2
18	52	24.4	131	4	US-08-765-012A-6
19	52	24.4	131	4	US-08-765-012A-7
20	52	24.4	131	4	US-08-765-012A-8
21	52	24.4	131	4	US-08-765-012A-9
22	52	24.4	131	4	US-08-765-012A-10
23	52	24.4	222	3	US-09-066-408-6
24	51	23.9	488	2	US-08-823-516-141
25	50	23.5	133	1	US-08-234-812-2
26	50	23.5	133	2	US-08-663-809-2
27	50	23.5	287	3	US-08-937-271-2

28	50	23.5	428	1	US-07-882-790-4	Sequence 4, Appli
29	50	23.5	613	2	US-08-915-207-2	Sequence 2, Appli
30	50	23.5	613	2	US-08-915-207-4	Sequence 4, Appli
31	50	23.5	613	4	US-09-238-555-2	Sequence 2, Appli
32	50	23.5	613	4	US-09-238-555-4	Sequence 4, Appli
33	50	23.5	1312	4	US-09-345-882-29	Sequence 29, Appli
34	49	23.0	139	4	US-08-559-896B-4	Sequence 4, Appli
35	49	23.0	451	4	US-09-134-001C-4461	Sequence 4461, Ap
36	49	23.0	652	4	US-08-559-896B-2	Sequence 2, Appli
37	49	23.0	1240	3	US-08-930-996A-4	Sequence 4, Appli
38	49	23.0	1388	4	US-09-572-191-2	Sequence 2, Appli
39	49	23.0	1388	4	US-09-723-262-2	Sequence 2, Appli
40	49	23.0	1388	4	US-09-723-219-2	Sequence 2, Appli
41	48.5	22.8	239	3	US-09-479-309-2	Sequence 2, Appli
42	48	22.5	46	1	US-08-214-770-11	Sequence 11, Appli
43	48	22.5	46	4	US-08-734-607B-20	Sequence 20, Appli
44	48	22.5	46	5	PCT-US95-02885-11	Sequence 11, Appli
45	48	22.5	223	3	US-09-066-408-7	Sequence 7, Appli

ALIGNMENTS

RESULT 1
US-09-058-489-27
; Sequence 27, Application US/09058489
; Patent No. 6103886
; GENERAL INFORMATION:
; APPLICANT: Whitehead Institute for Biomedical Research
; APPLICANT: Lahn, Bruce
; TITLE OF INVENTION: Genes in the No. 6103886-Recombining Region of
; FILE REFERENCE: WHI97-08pA
; CURRENT APPLICATION NUMBER: US/09/058,489
; CURRENT FILING DATE: 1998-04-10
; EARLIER APPLICATION NUMBER: 60/041,877
; EARLIER FILING DATE: 1997-04-11
; NUMBER OF SEQ ID NOS: 91
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 27
; LENGTH: 44
; TYPE: PRT
; ORGANISM: Human
US-09-058-489-27

Query Match	45.8%	Score	97.5	DB 3	Length	44
Best Local Similarity	59.0%	Pred. No.	4.2e-06			
Matches	23	Conservative	4	Mismatches	11	Indels 1; Gaps 1;
QY	1	MAHKLDLEIASLDRKAKLKA	TEMQ-KNTLMTKETTEQEK	38		
Db	1	MSDRPMAIEKFKDKLKKTKTQEKNP	LPSTETIQEK	39		

RESULT 2
US-08-664-856A-2
; Sequence 2, Application US/08664856A
; Patent No. 5663071
; GENERAL INFORMATION:
; APPLICANT: BRUCE R. ZETTER AND LERE BAO
; TITLE OF INVENTION: HUMAN THYMOSIN 15 GENE,
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN
; STREET: 130 WATER STREET
; CITY: BOSTON
; STATE: MA
; COUNTRY: USA
; ZIP: 02019
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSEQ Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/664,856A
FILING DATE: 17 JUN 1996
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: DAVID, RESNICK S
REGISTRATION NUMBER: 34,235
REFERENCE/DOCKET NUMBER: 46507
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-523-3400
TELEFAX: 617-523-6440
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 45 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: internal
ORIGINAL SOURCE:
US-08-664-856A-2

Query Match 44.8%; Score 95.5; DB 1; Length 45;
Best Local Similarity 53.8%; Pred. No. 7.9e-06;
Matches 21; Conservative 7; Mismatches 10; Indels 1; Gaps 1;

QY 1 MAHKLDLEETASLDKAKLKATEM-QKNTLMTKETTEQEK 38
| : | | | : : | | | | : | | | | : | | | : | | |
Db 1 MSKDPOLSEVETFDKSKLKKNTTEKNTLPKSTIQEKK 39

RESULT 3

US-08-801-796-2
Sequence 2, Application US/08801796
Patent No. 5721337
GENERAL INFORMATION:
APPLICANT: BRUCE R. ZETTER AND LERE BAO
TITLE OF INVENTION: HUMAN THYMOSIN 15 GENE,
PROTEIN AND USES THEREOF
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN
STREET: 130 WATER STREET
CITY: BOSTON
STATE: MA
COUNTRY: USA
ZIP: 02019
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSEQ Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/801,796
FILING DATE: 14-FEB-1997
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/664,856
FILING DATE: 17 JUN 1996
ATTORNEY/AGENT INFORMATION:
NAME: DAVID, RESNICK S
REGISTRATION NUMBER: 34,235
REFERENCE/DOCKET NUMBER: 46507
TELECOMMUNICATION INFORMATION:

TELEPHONE: 617-523-3400
TELEFAX: 617-523-6440
TELEX: 200291 STRE
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 45 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: internal
ORIGINAL SOURCE:
US-08-801-796-2

Query Match 44.8%; Score 95.5; DB 1; Length 45;
Best Local Similarity 53.8%; Pred. No. 7.9e-06;
Matches 21; Conservative 7; Mismatches 10; Indels 1; Gaps 1;

QY 1 MAHKLDLEETASLDKAKLKATEM-QKNTLMTKETTEQEK 38
| : | | | : : | | | | : | | | | : | | | : | | |
Db 1 MSKDPOLSEVETFDKSKLKKNTTEKNTLPKSTIQEKK 39

RESULT 4

US-08-931-877-2
Sequence 2, Application US/08931877
Patent No. 5831033
GENERAL INFORMATION:
APPLICANT: BRUCE R. ZETTER AND LERE BAO
TITLE OF INVENTION: HUMAN THYMOSIN 15 GENE,
PROTEIN AND USES THEREOF
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN
STREET: 130 WATER STREET
CITY: BOSTON
STATE: MA
COUNTRY: USA
ZIP: 02019
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSEQ Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/931,877
FILING DATE: 17-SEP-1997
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/801,796
FILING DATE: 14-FEB-1997
APPLICATION NUMBER: 08/664,856
FILING DATE: 17 JUN 1996
ATTORNEY/AGENT INFORMATION:
NAME: DAVID, RESNICK S
REGISTRATION NUMBER: 34,235
REFERENCE/DOCKET NUMBER: 46507
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-523-3400
TELEFAX: 617-523-6440
TELEX: 200291 STRE
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 45 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: internal

APPLICANT: WIEBEN, ERIC D.
TITLE OF INVENTION: STABLY-TRANSFORMED CELLS EXPRESSING
HUMAN THIOPURINE METHYLTRANSFERASE

NAME: DAVID, RESNICK S
REGISTRATION NUMBER: 34,235

NUMBER OF SEQUENCES: 18
CORRESPONDENCE ADDRESS:
ADDRESSEE: Schwegman, Lundberg & Woessner, P.A.
STREET: 3500 IDS Center
CITY: Minneapolis
STATE: MN
COUNTRY: USA
ZIP: 55402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/317,707
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/074,348
FILING DATE: 09-JUN-1993
ATTORNEY/AGENT INFORMATION:
NAME: Woessner, Warren D.
REGISTRATION NUMBER: 30,440
REFERENCE/DOCKET NUMBER: 150.103-US-01
TELECOMMUNICATION INFORMATION:
TELEPHONE: 612-339-3061
TELEFAX: 612-339-3061
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 245 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-317-707-2

Query Match 25.6%; Score 54.5; DB 1; Length 245;
Best Local Similarity 46.4%; Pred. No. 11;
Matches 13; Conservative 5; Mismatches 9; Indels 1; Gaps 1;

QY 12 SLDKAKLKATMOKNTLMTKETTEQKW 39
||| : ||:|||| :|| :|||
Db 7 SLDIEYSDEVQKNQVLTLEEW-QDKW 33

RESULT 11
US-08-514-921-2
Sequence 2, Application US/08514921
Patent No. 5856095
GENERAL INFORMATION:
APPLICANT: Evans, William E.
TITLE OF INVENTION: Identification of Two No. 5856095el Mutant
TITLE OF INVENTION: Alleles of Human Thiopurine S-Methyltransferase, and
TITLE OF INVENTION: Diagnostic Uses Thereof
NUMBER OF SEQUENCES: 38
CORRESPONDENCE ADDRESS:
ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
STREET: 1100 New York Avenue, N.W.
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/514,921
FILING DATE: 14-AUG-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Fleschner, Raz E.
REGISTRATION NUMBER: 34,331
REFERENCE/DOCKET NUMBER: 0656.0580000
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-371-2600
TELEFAX: 202-371-2540
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 245 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-514-921-4

Query Match 25.6%; Score 54.5; DB 2; Length 245;
Best Local Similarity 46.4%; Pred. No. 11;
Matches 13; Conservative 5; Mismatches 9; Indels 1; Gaps 1;

QY 12 SLDKAKLKATMOKNTLMTKETTEQKW 39
||| : ||:|||| :|| :|||
Db 7 SLDIEYSDEVQKNQVLTLEEW-QDKW 33

REGISTRATION NUMBER: 34,331
REFERENCE/DOCKET NUMBER: 0656.0580000
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-371-2600
TELEFAX: 202-371-2540
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 245 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-514-921-2

Query Match 25.6%; Score 54.5; DB 2; Length 245;
Best Local Similarity 46.4%; Pred. No. 11;
Matches 13; Conservative 5; Mismatches 9; Indels 1; Gaps 1;

QY 12 SLDKAKLKATMOKNTLMTKETTEQKW 39
||| : ||:|||| :|| :|||
Db 7 SLDIEYSDEVQKNQVLTLEEW-QDKW 33

RESULT 12
US-08-514-921-4
Sequence 4, Application US/08514921
Patent No. 5856095
GENERAL INFORMATION:
APPLICANT: Evans, William E.
TITLE OF INVENTION: Identification of Two No. 5856095el Mutant
TITLE OF INVENTION: Alleles of Human Thiopurine S-Methyltransferase, and
TITLE OF INVENTION: Diagnostic Uses Thereof
NUMBER OF SEQUENCES: 38
CORRESPONDENCE ADDRESS:
ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
STREET: 1100 New York Avenue, N.W.
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/514,921
FILING DATE: 14-AUG-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Fleschner, Raz E.
REGISTRATION NUMBER: 34,331
REFERENCE/DOCKET NUMBER: 0656.0580000
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-371-2600
TELEFAX: 202-371-2540
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 245 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-514-921-4

Query Match 25.6%; Score 54.5; DB 2; Length 245;
Best Local Similarity 46.4%; Pred. No. 11;
Matches 13; Conservative 5; Mismatches 9; Indels 1; Gaps 1;

QY 12 SLDKAKLKATMOKNTLMTKETTEQKW 39
||| : ||:|||| :|| :|||
Db 7 SLDIEYSDEVQKNQVLTLEEW-QDKW 33

```
RESULT 13
US-08-514-921-6
; Sequence 6, Application US/08514921
; Patent No. 5856095
; GENERAL INFORMATION:
; APPLICANT: Evans, William E.
; TITLE OF INVENTION: Identification of Two No. 5856095el Mutant
; TITLE OF INVENTION: Alleles of Human Thiopurine S-Methyltransferase, and
; TITLE OF INVENTION: Diagnostic Uses Thereof
; NUMBER OF SEQUENCES: 38
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
; STREET: 1100 New York Avenue, N.W.
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/514.921
; FILING DATE: 14-AUG-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Fleshner, Raz E.
; REGISTRATION NUMBER: 34,331
; REFERENCE/DOCKET NUMBER: 0656.0580000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-371-2540
; TELEFAX: 202-371-2540
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 245 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-514-921-6

Query Match 25.6%; Score 54.5; DB 2; Length 245;
Best Local Similarity 46.4%; Pred. No. 11;
Matches 13; Conservative 5; Mismatches 9; Indels 1; Gaps 1;

QY 12 SLDKAKLKATMOKMTLMTKETTEQEKW 39
Db 7 SLDIEYSDEVQKNQVLTLEW-QDKW 33

RESULT 14
US-08-514-921-8
; Sequence 6, Application US/08514921
; Patent No. 5856095
; GENERAL INFORMATION:
; APPLICANT: Evans, William E.
; TITLE OF INVENTION: Identification of Two No. 5856095el Mutant
; TITLE OF INVENTION: Alleles of Human Thiopurine S-Methyltransferase, and
; TITLE OF INVENTION: Diagnostic Uses Thereof
; NUMBER OF SEQUENCES: 38
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
; STREET: 1100 New York Avenue, N.W.
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/514.921
; FILING DATE: 14-AUG-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Fleshner, Raz E.
; REGISTRATION NUMBER: 34,331
; REFERENCE/DOCKET NUMBER: 0656.0580000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-371-2540
; TELEFAX: 202-371-2540
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 245 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-514-921-8

Query Match 25.6%; Score 54.5; DB 2; Length 245;
Best Local Similarity 46.4%; Pred. No. 11;
Matches 13; Conservative 5; Mismatches 9; Indels 1; Gaps 1;

QY 12 SLDKAKLKATMOKMTLMTKETTEQEKW 39
Db 7 SLDIEYSDEVQKNQVLTLEW-QDKW 33

RESULT 15
US-08-858-207A-269
; Sequence 269, Application US/08858207A
; Patent No. 6348328
; GENERAL INFORMATION:
; APPLICANT: Black, Michael
; APPLICANT: Hodgson, John
; APPLICANT: Knowles, David
; APPLICANT: Nicholas, Richard
; APPLICANT: Stodola, Robert
; TITLE OF INVENTION: No. 6348328el Compounds
; NUMBER OF SEQUENCES: 552
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19406-0939
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/858,207A
; FILING DATE: 09-MAY-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/017670
; FILING DATE: 14-MAY-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Gimmi, Edward R
; REGISTRATION NUMBER: 38,891
; REFERENCE/DOCKET NUMBER: P50475
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-270-4478
; TELEFAX: 610-270-5090
; TELEX:
; INFORMATION FOR SEQ ID NO: 269:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 210 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
```

; TOPOLOGY: linear
; MOLECULE TYPE: No. 6348328e
US-08-858-207A-269

Query Match 25.1%; Score 53.5; DB 4; Length 210;
Best Local Similarity 30.0%; Pred. No. 12;
Matches 15; Conservative 9; Mismatches 7; Indels 19; Gaps 2;

Oy 5 LDLEEIASLDKAKLKATEMOK-----NTLMTKETTEQEKWSE 41
|||::| || ||:: :||:| |:
Db 90 LLDLKILPD-----TELEETKNYIKSQIISDLDVLTSEKSEAKHWE 133

Search completed: May 30, 2003, 10:36:05
Job time : 27 secs

THIS PAGE BLANK (1/1)


```

AC Q9D9W8;
DT 01-JUN-2001 (Tremblrel. 17, Created)
DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)
DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)
DE Adult male testis cDNA, RIKEN full-length enriched library,
DE clone:1700026G08, full insert sequence.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=TESTIS;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gliss C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bull C., Fletcher C., Fujita M., Gariboldi M.,
RA Gusticich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
DR EMBL; AK006383; BAB24560.1; -
DR InterPro; IPR001152; Thymosin_b4.
DR Pfam; PF01290; Thymosin; 1.
DR ProDom; PD005116; Thymosin_b4; 1.
DR SMART; SM00152; THY; 1.
DR PROSITE; PS00500; THYMOSIN_B4; 1.
SQ SEQUENCE 38 AA; 4418 MW; 42BCD4A0E30ED5AA CRC64;

Query Match 51.4%; Score 109.5; DB 11; Length 38;
Best Local Similarity 75.0%; Pred. No. 5,1e-06;
Matches 27; Conservative 1; Mismatches 7; Indels 1; Gaps 1;

QY 9 EIASLDRKLKATEMQ-KNTLMTKETTEQKWEIS 43
| | | | | | | | | | | | | | | | | | | | | |
DB 3 EIASFHKAKLKTETQKNTLPTRETIEQKRSEIS 38
| | | | | | | | | | | | | | | | | | | | | |

RESULT 3
Q76538 ID 076538 PRELIMINARY; PRT; 41 AA.
AC 076538;
DT 01-NOV-1998 (Tremblrel. 08, Created)
DT 01-NOV-1998 (Tremblrel. 08, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE Thymosin beta.
OS Strongylocentrotus purpuratus (Purple sea urchin).
OC Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
OC Echinoidea; Euechinoidea; Echinacea; Echinoidea; Strongylocentrotidae;
OC Strongylocentrotus.
OX NCBI_TaxID=7668;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=99328904; PubMed=10398804;
RA Pancer Z., Rast J.P., Davidson E.H.;
RT "Origins of immunity: transcription factors and homologues of effector
RT genes of the vertebrate immune system expressed in sea urchin
RT coelomocytes.";
RL Immunogenetics 49:773-786(1999).
DR EMBL; AF076515; AAC26833.1; -

```

```

DR InterPro; IPR001152; Thymosin_b4.
DR Pfam; PF01290; Thymosin; 1.
DR SMART; SM00152; THY; 1.
DR PROSITE; PS00500; THYMOSIN_B4; UNKNOWN1.
SQ SEQUENCE 41 AA; 4598 MW; B4A7838CE16915B8 CRC64;

Query Match 49.5%; Score 105.5; DB 5; Length 41;
Best Local Similarity 61.5%; Pred. No. 1.6e-05;
Matches 24; Conservative 4; Mismatches 10; Indels 1; Gaps 1;

QY 1 MAHKLDLEETASLDKAKLKATEM-QKNTLMTKETTEQEK 38
| | | | | | | | | | | | | | | | | | | | | |
DB 1 MADKPDVSAVSSFDKTKLKKTETEEKNTLPTKETIEQEK 39
| | | | | | | | | | | | | | | | | | | | | |

RESULT 4
Q9GUA6 ID 09GUA6 PRELIMINARY; PRT; 42 AA.
AC 09GUA6;
DT 01-MAR-2001 (Tremblrel. 16, Created)
DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE Beta-thymosin.
OS Sycon raphanus.
OC Eukaryota; Metazoa; Porifera; Calcarea; Calcaronea; Leucosoleniida;
OC Sycettidae.
OX NCBI_TaxID=56443;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=20496958; PubMed=11040289;
RA Manuel M., Kruse M., Muller W.E., Le Parco Y.;
RT "The comparison of beta-thymosin homologues Among Metazoa Supports an
RT Arthropod-Nematode Clade.";
RL J. Mol. Evol. 51:378-381(2000).
DR EMBL; AF155935; AAG08963.1; -
DR InterPro; IPR001152; Thymosin_b4.
DR Pfam; PF01290; Thymosin; 1.
DR SMART; SM00152; THY; 1.
SQ SEQUENCE 42 AA; 4637 MW; BD9BF6EA8CE16C12 CRC64;

Query Match 49.1%; Score 104.5; DB 5; Length 42;
Best Local Similarity 61.5%; Pred. No. 2.2e-05;
Matches 24; Conservative 3; Mismatches 11; Indels 1; Gaps 1;

QY 1 MAHKLDLEETASLDKAKLKATE-MQKNTLMTKETTEQEK 38
| | | | | | | | | | | | | | | | | | | | | |
DB 1 MGDKPDVSEVAFDKTKLKKTETAEKNPLPTKETIEQEK 39
| | | | | | | | | | | | | | | | | | | | | |

RESULT 5
Q9GUA5 ID 09GUA5 PRELIMINARY; PRT; 42 AA.
AC 09GUA5;
DT 01-MAR-2001 (Tremblrel. 16, Created)
DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE Beta-thymosin.
OS Sycon raphanus.
OC Eukaryota; Metazoa; Porifera; Calcarea; Calcaronea; Leucosoleniida;
OC Sycettidae.
OX NCBI_TaxID=56443;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=20496958; PubMed=11040289;
RA Manuel M., Kruse M., Muller W.E., Le Parco Y.;
RT "The comparison of beta-thymosin homologues Among Metazoa Supports an
RT Arthropod-Nematode Clade.";
RL J. Mol. Evol. 51:378-381(2000).
DR EMBL; AF156177; AAG08964.1; -
DR InterPro; IPR001152; Thymosin_b4.
DR Pfam; PF01290; Thymosin; 1.
DR SMART; SM00152; THY; 1.
SQ SEQUENCE 42 AA; 4706 MW; BD9BF75A8CE16C12 CRC64;

```

```
Query Match 49.1%; Score 104.5; DB 5; Length 42;
Best Local Similarity 61.5%; Pred. No. 2.2e-05;
Matches 24; Conservative 3; Mismatches 11; Indels 1; Gaps 1;

QY 1 MAHKLDLEETIASLDKAKLKATE-MQKNTLMTKTTTQEK 38
  | | | | | | | | | | | | | | | | | | | | | |
DB 1 MGDKPDVSEVAFDKTKLAKTETAEKNPLPTKTETQEK 39
  | | | | | | | | | | | | | | | | | | | | | |

RESULT 6
Q9NQ05 PRELIMINARY; PRT; 44 AA.
AC Q9NQ05;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE DJ1071110.1 (novel thymosin/interferon-inducible multigene family
  member).
GN DJ1071110.1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Tromans A.;
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
RL EMBL: AL133228; CAB94229.1; -
DR InterPro; IPR001152; Thymosin_b4.
DR Pfam; PF01290; Thymosin; 1.
DR ProDom; PD005116; Thymosin_b4; 1.
DR SMART; SM00152; THY; 1.
DR PROSITE; PS00500; THYMOSIN_B4; 1.
SQ SEQUENCE 44 AA; 5043 MW; D7C35E58482DAAD2 CRC64;

Query Match 46.2%; Score 98.5; DB 4; Length 44;
Best Local Similarity 59.0%; Pred. No. 0.00012;
Matches 23; Conservative 4; Mismatches 11; Indels 1; Gaps 1;

QY 1 MAHKLDLEETIASLDKAKLKATE-MQKNTLMTKTTTQEK 38
  | | | | | | | | | | | | | | | | | | | | | |
DB 1 MSDKSDMAEIEFKDKSLKAKTETQENPLPSKETIQEK 39
  | | | | | | | | | | | | | | | | | | | | | |

RESULT 7
Q95274 PRELIMINARY; PRT; 46 AA.
AC Q95274;
DT 01-FEB-1997 (TREMBLrel. 02, Created)
DT 01-FEB-1997 (TREMBLrel. 02, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE Thymosin beta-4 (Fragment).
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=SMALL INTESTINE;
RA Winteroe A.K., Fredholm M., Davies W.;
RT "Evaluation and characterization of a porcine small intestine cDNA
  library."
RL Submitted (JUL-1995) to the EMBL/GenBank/DBJ databases.
RL EMBL: Z81195; CAB03562.1; -
DR InterPro; IPR001152; Thymosin_b4.
DR Pfam; PF01290; Thymosin; 1.
DR ProDom; PD005116; Thymosin_b4; 1.
DR SMART; SM00152; THY; 1.
DR PROSITE; PS00500; THYMOSIN_B4; 1.
FT NON_TER 1
SQ SEQUENCE 46 AA; 5248 MW; 347C6821EBEFAA94 CRC64;
```

```
Query Match 45.3%; Score 96.5; DB 6; Length 46;
Best Local Similarity 59.0%; Pred. No. 0.00021;
Matches 23; Conservative 3; Mismatches 12; Indels 1; Gaps 1;

QY 1 MAHKLDLEETIASLDKAKLKATE-MQKNTLMTKTTTQEK 38
  | | | | | | | | | | | | | | | | | | | | | |
DB 3 MXDKPDMAEIEFKDKSLKAKTETQENPLPSKETIQEK 41
  | | | | | | | | | | | | | | | | | | | | | |

RESULT 8
Q962B8 PRELIMINARY; PRT; 45 AA.
AC Q962B8;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Thymosin beta-4.
OS Branchiostoma belcheri (Amphioxus).
OC Eukaryota; Metazoa; Chordata; Cephalochordata; Branchiostomidae;
OC Branchiostoma.
OX NCBI_TaxID=7741;
RN [1]
RP SEQUENCE FROM N.A.
RA Huang X., Han H., Yang H., Zhang H.;
RT "Cloning of thymosin beta-4 gene in amphioxus."
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
RL EMBL: AY037883; AAK72482.1; -
DR InterPro; IPR001152; Thymosin_b4.
DR Pfam; PF01290; Thymosin; 1.
DR ProDom; PD005116; Thymosin_b4; 1.
DR SMART; SM00152; THY; 1.
SQ SEQUENCE 45 AA; 5180 MW; 96E51C7852C37D7F CRC64;

Query Match 44.8%; Score 95.5; DB 5; Length 45;
Best Local Similarity 56.4%; Pred. No. 0.00027;
Matches 22; Conservative 4; Mismatches 12; Indels 1; Gaps 1;

QY 1 MAHKLDLEETIASLDKAKLKATE-MQKNTLMTKTTTQEK 38
  | | | | | | | | | | | | | | | | | | | | | |
DB 1 MSDKPDLSVEVFKDKSLKAKTETSVKNTLPSKEAIEQEK 39
  | | | | | | | | | | | | | | | | | | | | | |

RESULT 9
P97563 PRELIMINARY; PRT; 45 AA.
AC P97563;
DT 01-MAY-1997 (TREMBLrel. 03, Created)
DT 01-MAY-1997 (TREMBLrel. 03, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Thymosin beta-like protein.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=DUNNING R-3327 PROSTATIC ADENOCARCINOMA;
RA Bao L., Zetter B.R.;
RL Submitted (APR-1995) to the EMBL/GenBank/DBJ databases.
RL EMBL: U25684; AAB37101.1; -
DR InterPro; IPR001152; Thymosin_b4.
DR Pfam; PF01290; Thymosin; 1.
DR SMART; SM00152; THY; 1.
DR PROSITE; PS00500; THYMOSIN_B4; UNKNOWN.1.
SQ SEQUENCE 45 AA; 5304 MW; 653E304556ED4153 CRC64;

Query Match 44.8%; Score 95.5; DB 11; Length 45;
Best Local Similarity 53.8%; Pred. No. 0.00027;
Matches 21; Conservative 7; Mismatches 10; Indels 1; Gaps 1;

QY 1 MAHKLDLEETIASLDKAKLKATE-MQKNTLMTKTTTQEK 38
  | | | | | | | | | | | | | | | | | | | | | |
DB 1 MSDKPDLSVEVFKDKSLKAKTETSVKNTLPSKETIQEK 39
  | | | | | | | | | | | | | | | | | | | | | |
```


THIS PAGE BLANK (1-3PTO)

GenCore version 5.1.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 30, 2003, 10:31:31 ; Search time 21 Seconds
(without alignments)
84.928 Million cell updates/sec

Title: US-09-915-178-2

Perfect score: 213

Sequence: 1 MAHKLDLEIASLDKAKLA.....QKNTLMKTETQEKWSEIS 43

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	130.5	61.3	43	1 TYBO_HUMAN	P13472 homo sapien
2	112.5	52.8	41	1 TYB9_BOVIN	P21752 bos taurus
3	110.5	51.9	41	1 TYB9_PIG	P21753 sus scrofa
4	108.5	50.9	42	1 TYB8_CYPCA	Q91954 cyprinus ca
5	106.5	50.0	41	1 TYBA_ONCMY	P26351 oncorhynch
6	106.5	50.0	43	1 TYBO_TORMA	Q91980 torpedo mar
7	99.5	46.7	42	1 TYBB_ONCMY	P26352 oncorhynch
8	97.5	45.8	50	1 TYB4_MOUSE	P20065 mus musculu
9	96.5	45.3	43	1 TYBB_LATJA	P33248 lateolabrax
10	95.5	44.8	43	1 TYB4_RABIT	P34032 coryctolagus
11	95.5	44.8	43	1 TYB4_XENLA	P18758 xenopus lae
12	95.5	44.8	44	1 TYB4_BRARE	Q9W7M8 brachydanio
13	92.5	43.4	43	1 TYB4_HUMAN	P01253 homo sapien
14	90.5	42.5	44	1 TYBN_COTJA	Q9det5 coturnix co
15	88.5	40.6	44	1 TYBN_HUMAN	Q99406 homo sapien
16	85.5	40.1	43	1 TYB_GILMI	Q9dfj9 gillichthys
17	83.5	39.2	43	1 TYBY_HUMAN	O14604 homo sapien
18	70.5	33.1	45	1 TYBA_CYPCA	Q91955 cyprinus ca
19	63	29.6	543	1 HMA1_ARATH	P42804 arabidopsis
20	60	28.2	2869	1 RBP1_PLAVB	Q00798 plasmodium
21	58	27.2	418	1 HEM1_ECOLI	P13580 escherichia
22	57	26.8	467	1 P60_LISIN	Q01836 listeria in
23	56.5	26.5	161	1 Y0UD_CAEEL	P30631 caenorhabdi
24	56	26.3	327	1 ODBB_BACSU	P37941 bacillus su
25	56	26.3	3396	1 PGCV_HUMAN	P13611 homo sapien
26	55.5	26.1	757	1 PPE2_MOUSE	Q35385 mus musculu
27	54.5	25.6	245	1 TPMT_MOUSE	P51580 homo sapien
28	54.5	25.6	432	1 ACDB_HUMAN	P45954 h acyl-coa
29	54.5	25.6	1433	1 REST_CHICK	Q02184 gallus gall
30	54.5	25.6	4473	1 PLE1_CRIGR	Q9155 cricetus
31	54.5	25.6	4687	1 PLE1_RAT	P30427 rattus norv
32	54	25.4	742	1 KM65_YEAST	Q03656 saccharomyc
33	54	25.4	2663	1 CENE_HUMAN	Q02224 homo sapien

ALIGNMENTS

RESULT 1

ID	TYBO_HUMAN	STANDARD;	PRT;	43 AA.
AC	P13472;			
DT	01-JAN-1990 (Rel. 13, Created)			
DT	01-JAN-1990 (Rel. 13, Last sequence update)			
DT	15-JUN-2002 (Rel. 41, Last annotation update)			
DE	Thymosin beta-10.			
GN	TMSB10 OR PTMB10 OR THYB10.			
OS	Homo sapiens (Human).			
'OS	Rattus norvegicus (Rat), and			
OC	Equus caballus (Horse).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
OX	NCBI_TaxID=9606, 10116, 9796;			
[1]	SEQUENCE FROM N.A.			
RP	SPECIES=Human; TISSUE=Kidney;			
RC	MEDLINE=88209071; PubMed=33655256;			
RA	McCreary V., Kartha S., Bell G.I., Toback F.G.;			
RT	"Sequence of a human kidney cDNA clone encoding thymosin beta 10.";			
RL	Biochem. Biophys. Res. Commun. 152:862-866(1988).			
[2]	SEQUENCE FROM N.A.			
RP	SPECIES=Human; TISSUE=Brain;			
RC	MEDLINE=90384336; PubMed=2169566;			
RA	Hall A.K., Hempstead J., Morgan J.I.;			
RT	"Thymosin beta 10 levels in developing human brain and its regulation			
RL	by retinoic acid in the HTB-10 neuroblastoma.";			
[3]	Brain Res. Mol. Brain Res. 8:129-135(1990).			
RP	SEQUENCE FROM N.A.			
RC	SPECIES=Human;			
RA	Weterman M.A., van Muijen G.N., Ruiters D.J., Bloemers H.P.;			
RT	"Thymosin beta-10 expression in melanoma cell lines and melanocytic			
RL	lesions: a new progression marker for human cutaneous melanoma.";			
[4]	Int. J. Cancer 53:278-284(1993).			
RP	SEQUENCE FROM N.A.			
RC	SPECIES=Human;			
RA	Condon M.R., Hall A.K.;			
RT	Submitted (APR-1992) to the EMBL/GenBank/DBJ databases.			
[5]	SEQUENCE FROM N.A.			
RP	SPECIES=Human; TISSUE=Brain, and Uterus;			
RC	Strausberg R.;			
RA	Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.			
[6]	SEQUENCE FROM N.A.			
RP	SPECIES=Human;			
RC	Goodall G.J., Horecker B.L.;			
RA	"Molecular cloning of the cDNA for rat spleen thymosin beta 10 and			
RT	the deduced amino acid sequence.";			
RL	Arch. Biochem. Biophys. 256:402-405(1987).			

34	53.5	25.1	178	1 SIGY_BACSU	P94370 bacillus su
35	53.5	25.1	816	1 NEL2_MOUSE	O61220 mus musculu
36	53.5	25.1	1005	1 Y321_MYCPN	P75327 mycoplasma
37	53.5	25.1	4684	1 PLE1_HUMAN	Q15149 homo sapien
38	53	24.9	220	1 Y231_RICPR	Q92du3 rickettsia
39	53	24.9	863	1 GLND_HABIN	P43919 haemophilus
40	53	24.9	1107	1 POL2_RRVS	P36324 raspberry r
41	52.5	24.6	205	1 ALAG_RAT	P02764 rattus norv
42	52.5	24.6	509	1 RA18_MOUSE	Q9qkx2 mus musculu
43	52	24.4	120	1 ACPS_STRPN	Q9f75 streptococc
44	52	24.4	222	1 CAS2_BOVIN	P02663 bos taurus
45	52	24.4	398	1 SUCC_BRUME	O8Yj66 brucella me


```

DR PROSITE; PS00500; THYMOSIN_B4; 1.
KW Actin-binding; Cytoskeleton; Acetylation; 3D-structure.
FT INIT_MET 0
FT PEPTIDE 1 41 THYMOSIN BETA-9.
FT PEPTIDE 1 39 THYMOSIN BETA-8.
FT MOD_RES 1 1 ACETYLATION.
SQ SEQUENCE 41 AA; 4674 MW; 268E1AB71E6A8909 CRC64;

Query Match
Best Local Similarity 52.8%; Score 112.5; DB 1; Length 41;
Matches 28; Conservative 0; Mismatches 9; Indels 1; Gaps 1;

Qy 2 AHKLDLEIASLDKAKLKATQM-Q-KNTLMKTEQEK 38
Db 1 ADKPDGGEINSPDKAKLKATQKNTLPTKETIEQEK 38

RESULT 3
TYB9_PIG
ID TYB9_PIG STANDARD; PRT; 41 AA.
AC P21753;
DT 01-MAY-1991 (Rel. 18, Created)
DT 01-MAY-1991 (Rel. 18, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Thymosin beta-9.
OS Sus scrofa (Fig).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE.
RC TISSUE=Thymus;
RX MEDLINE=89372897; PubMed=2774558;
RA Hannappel E., Wartenberg F., Bustelo X.R.;
RT "Isolation and characterization of thymosin beta 9 Met from pork
  spleen.";
RL Arch. Biochem. Biophys. 273:396-402(1989).
RN [2]
RP SEQUENCE.
RC TISSUE=Thymus;
RX MEDLINE=91216695; PubMed=2090639;
RA Low T.L.K., Lin C.Y., Pan T.L., Chiou A.J., Tsugita A.;
RT "Structure and immunological properties of thymosin beta 9 Met, a new
  analog of thymosin beta 4 isolated from porcine thymus.";
RL Int. J. Pept. Protein Res. 36:481-488(1990).
CC FUNCTION: Plays an important role in the organization of the
  cytoskeleton. Binds to and sequesters actin monomers (G actin) and
  therefore inhibits actin polymerization (By similarity).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC PIR; S05644; S05644.
CC PIR; A60290; A60290.
DR HSSP; P21752; IHJ0.
DR InterPro; IPR001152; Thymosin_b4.
DR Pfam; PF01290; Thymosin; 1.
DR ProDom; PD005116; Thymosin_b4; 1.
DR SMART; SM00152; THY; 1.
DR PROSITE; PS00500; THYMOSIN_B4; 1.
KW Actin-binding; Cytoskeleton; Acetylation.
FT MOD_RES 1 1
SQ SEQUENCE 41 AA; 4692 MW; 268E1AB6B5703909 CRC64;

Query Match
Best Local Similarity 51.9%; Score 110.5; DB 1; Length 41;
Matches 27; Conservative 1; Mismatches 9; Indels 1; Gaps 1;

Qy 2 AHKLDLEIASLDKAKLKATQM-Q-KNTLMKTEQEK 38
Db 1 ADKPDGGEINSPDKAKLKATQKNTLPTKETIEQEK 38

RESULT 4
TYBB_CYPCA

```

```

ID TYBB_CYPCA STANDARD; PRT; 42 AA.
QY1954;
AC Q91954;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Thymosin beta-b.
OS Cyprinus carpio (Common carp).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Cyprinus.
OX NCBI_TaxID=7962;
RN [1]
RP SEQUENCE FROM N.A.
RA Fujiki K., Nakao M., Shin D., Yano T.;
RT "Molecular cloning of carp (Cyprinus carpio) thymosin beta b.";
RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
CC FUNCTION: Plays an important role in the organization of the
  cytoskeleton. Binds to and sequesters actin monomers (G actin) and
  therefore inhibits actin polymerization (By similarity).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: BELONGS TO THE THYMOSIN BETA FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
  between the Swiss Institute of Bioinformatics and the EMBL outstation
  at the European Bioinformatics Institute. There are no restrictions on its
  use by non-profit institutions as long as its content is in no way
  modified and this statement is not removed. Usage by and for commercial
  entities requires a license agreement (See http://www.isb-sib.ch/announce/
  or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AB028457; BAA96493.1; --
CC InterPro; IPR001152; Thymosin_b4.
CC Pfam; PF01290; Thymosin; 1.
CC SMART; SM00152; THY; 1.
CC PROSITE; PS00500; THYMOSIN_B4; 1.
KW Actin-binding; Cytoskeleton.
FT INIT_MET 0
SQ SEQUENCE 42 AA; 4838 MW; 47667BB96C9DDAFB CRC64;

Query Match
Best Local Similarity 50.9%; Score 108.5; DB 1; Length 42;
Matches 25; Conservative 3; Mismatches 12; Indels 1; Gaps 1;

Qy 2 AHKLDLEIASLDKAKLKATQM-Q-KNTLMKTEQEKWSE 41
Db 1 ADKPDISEVSQFDKTKLKKTKETQKNTLPTKETIEQKCE 41

RESULT 5
TYBA_ONCMY
ID TYBA_ONCMY STANDARD; PRT; 41 AA.
AC P26351; Q9PT32;
DT 01-MAY-1992 (Rel. 22, Created)
DT 01-MAY-1992 (Rel. 22, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Thymosin beta-11.
OS Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
OX NCBI_TaxID=8022;
RN [1]
RP SEQUENCE.
RC TISSUE=Spleen;
RX MEDLINE=92245862; PubMed=1575682;
RA Yialouris P.P., Coles B., Tsitsiloni O., Schmid B., Howell S.;
RT "The complete sequences of trout (Salmo gairdneri) thymosin beta 11
  and its homologue thymosin beta 12.";
RL Biochem. J. 283:385-389(1992).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;

```

modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).

EMBL; AJ276369; CAB76965.1; -
InterPro: IPR001152; Thymosin_b4.
Pfam: PF01290; Thymosin_1.
ProDom: PD005116; Thymosin_b4; 1.
SMART: SM00152; THY; 1.
PROSITE: PS00500; THYMOSIN_B4; 1.
KW Actin-binding; Cytoskeleton.
FT INIT_MET 0
SQ SEQUENCE 43 AA; 4785 MW; 63D5D49977A2E351 CRC64;

Query Match 50.0%; Score 106.5; DB 1; Length 43;
Best Local Similarity 58.1%; Pred. No. 4.2e-06;
Matches 25; Conservative 6; Mismatches 11; Indels 1; Gaps 1;

QY 2 AHKLDLEIASLDRKAKLKATEMQ-KNTLTMTKETQEOKWSEIS 43
||| ||| ||| ||| : ||| ||| ||| ||| : |||
Db 1 ADKPDFGEVASFDSKLKKTDTEVKNTLPFKETIDQEKKAESS 43

RESULT 7

TYBB_ONCMY STANDARD; PRT; 42 AA.

AC P26352;
DT 01-MAY-1992 (Rel. 22, Created)
DT 01-MAY-1992 (Rel. 22, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Thymosin beta-12.
OS Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
OX NCBI_TaxID=8022;
RN [1]
RP SEQUENCE.
RC TISSUE=Spleen;
RX MEDLINE=92246862; PubMed=1575682;
RA Yialouris P.P., Coles B., Tsitsiloni O., Schmid B., Howell S.,
RA Aikten A., Voelter W., Haritos A.A.;
RT "The complete sequences of trout (Salmo gairdneri) thymosin beta 11
and its homologue thymosin beta 12.";
RL Biochem. J. 283:385-389(1992).
RN [2]

RP SYNTHESIS.
RA Echner H., Yialouris P.P., Haritos A.A., Gruebler G., Voelter W.;
RT "Structure and syntheses of thymosin beta-11 and beta-12.";
RL (In) Schneider C.H., Eberles A.N. (eds.);
RL Peptides 1992, pp.751-752, Escom Science Publishers, Leiden (1993).
CC -!- FUNCTION: Plays an important role in the organization of the
CC cytoskeleton. Binds to and sequesters actin monomers (G actin) and
CC therefore inhibits actin polymerization (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- SIMILARITY: BELONGS TO THE THYMOSIN BETA FAMILY.
CC HSPSP; P21752; IHJ0.
DR InterPro: IPR001152; Thymosin_b4.
DR Pfam; PF01290; Thymosin_1.
DR SMART; SM00152; THY; 1.
DR PROSITE; PS00500; THYMOSIN_B4; 1.
KW Actin-binding; Cytoskeleton; Acetylation.
FT MOD_RES 1 1
FT ACETYLAION.
SQ SEQUENCE 42 AA; 4760 MW; C290EE1BF604CA90 CRC64;

Query Match 46.7%; Score 99.5; DB 1; Length 42;
Best Local Similarity 63.9%; Pred. No. 2.6e-05;
Matches 23; Conservative 3; Mismatches 9; Indels 1; Gaps 1;

QY 4 KLDLEIASLDRKAKLKATEMQ-KNTLTMTKETQEOK 38
||||| : ||| ||| ||| ||| ||| ||| ||| |||
Db 3 KPDLAEVSFDPDKLKKTKTQEKNPLPTKETIQEOK 38

```
RESULT 8
TYB4_MOUSE
ID TYB4_MOUSE STANDARD; PRT; 50 AA.
AC P20065;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE THYMOSIN beta-4 (T beta 4).
GN TMSB4 OR TMSB4X OR PTMB4.
OS Mus musculus (Mouse)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A. (LONG FORM).
RX MEDLINE=90278100; PubMed=2351831;
RA Rudin C.M., Engler P., Storb U.;
RT "Differential splicing of thymosin beta 4 mRNA.";
RL J. Immunol. 144:4857-4862(1990).
RN [2]
RP SEQUENCE FROM N.A. (SHORT FORM).
RX MEDLINE=91032153; PubMed=2226839;
RA Low T.L.K., Pan T.L., Lin Y.S.;
RT "Depression of prothymosin alpha production in murine thymus
correlates with staphylococcal enterotoxin-B-induced
immunosuppression.";
RL FEBS Lett. 273:1-5(1990).
RN [3]
RP SEQUENCE FROM N.A. (SHORT FORM).
RX STRAIN=129/SV;
RC MEDLINE=96435916; PubMed=8838802;
RA Li X., Zimmerman A., Copeland N.G., Gilbert D.J., Jenkins N.A.,
RA Yin H.L.;
RT "The mouse thymosin beta 4 gene: structure, promoter identification,
and chromosome localization.";
RL Genomics 32:388-394(1996).
CC -!- FUNCTION: Plays an important role in the organization of the
cytoskeleton. Binds to and sequesters actin monomers (G actin) and
therefore inhibits actin polymerization.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- ALTERNATIVE PRODUCTS: 2 isoforms; a long form (shown here) and a
short form; are produced by alternative splicing.
CC -!- TISSUE SPECIFICITY: ORIGINALLY FOUND IN THYMUS BUT IT IS WIDELY
DISTRIBUTED IN MANY TISSUES.
CC -!- SIMILARITY: BELONGS TO THE THYMOSIN BETA FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC
CC -----
DR EMBL; X16053; CAA34187.1; -
DR EMBL; X16053; CAA34188.1; -
DR EMBL; U38967; AAC52490.1; -
DR PIR; S08074; S08074.
DR PIR; S12884; S12884.
DR PIR; A37217; A37217.
DR MGD; P21752; IHJO.
DR GSD; MGI:99510; Tmsb4x.
DR InterPro; IPR001152; Thymosin_b4.
DR Pfam; PF01290; Thymosin; 1.
DR ProDom; PD005116; Thymosin_b4; 1.
DR SMART; SM00152; THY; 1.
DR PROSITE; PS00500; THYMOSIN_B4; 1.
DR ACTIN-binding; Cytoskeleton; Acetylation; Alternative splicing.
KW CHAIN 1 50 THYMOSIN BETA-4, LONG FORM.
FT CHAIN 8 50 THYMOSIN BETA-4, SHORT FORM.
FT MOD_RES 8 8 ACETYLATION (BY SIMILARITY).
FT VARSPLIC 1 6 MISSING (IN SHORT ISOFORM).
```

```
SQ SEQUENCE 50 AA; 5679 MW; 9A289F60EE48EB8A CRC64;
Query Match 45.8%; Score 97.5; DB 1; Length 50;
Best Local Similarity 59.0%; Pred. No. 5.2e-05;
Matches 23; Conservative 4; Mismatches 11; Indels 1; Gaps 1;

QY 1 MAHKDLEETASLDKAKLKATEMQ-KNTLMTKETTEQEK 38
I: | | | | | | | | | | | | | | | | | | | |
|: | | | | | | | | | | | | | | | | | | | |
7 MSKPDMAEIEKFDKSLKKTETQEKNPSPSKETIEQEK 45

Db
RESULT 9
TYBB_LATJA
ID TYBB_LATJA STANDARD; PRT; 43 AA.
AC P33248;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Thymosin beta-12.
OS Lateolabrax japonicus (Japanese sea perch) (Japanese sea bass).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Percoidae;
OC Moronidae; Lateolabrax.
OX NCBI_TaxID=8164;
RN [1]
RP SEQUENCE.
RC TISSUE=Liver;
RX MEDLINE=92117698; PubMed=1731637;
RA Low T.L.K., Liu D.T.-W., Jou J.-H.;
RT "Primary structure of thymosin beta 12, a new member of the beta-
thymosin family isolated from perch liver.";
RL Arch. Biochem. Biophys. 293:32-39(1992).
CC -!- FUNCTION: Plays an important role in the organization of the
cytoskeleton. Binds to and sequesters actin monomers (G actin) and
therefore inhibits actin polymerization (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- SIMILARITY: BELONGS TO THE THYMOSIN BETA FAMILY.
CC PIR; S21694; S21694.
DR HSSP; P21752; IHJO.
DR InterPro; IPR001152; Thymosin_b4.
DR Pfam; PF01290; Thymosin; 1.
DR SMART; SM00152; THY; 1.
DR PROSITE; PS00500; THYMOSIN_B4; 1.
KW ACTIN-binding; Cytoskeleton; Acetylation.
FT MOD_RES 1 1
FT MOD_RES 1 1 ACETYLATION.
SQ SEQUENCE 43 AA; 4779 MW; E7E4A32B091046DA CRC64;
Query Match 45.3%; Score 96.5; DB 1; Length 43;
Best Local Similarity 61.1%; Pred. No. 5.8e-05;
Matches 22; Conservative 3; Mismatches 10; Indels 1; Gaps 1;

QY 4 KLDLEETASLDKAKLKATEMQ-KNTLMTKETTEQEK 38
I: | | | | | | | | | | | | | | | | | | | |
|: | | | | | | | | | | | | | | | | | | | |
3 KPDISEVTSPDKTLKKTETQEKNPSPSKETIEQEK 38

Db
RESULT 10
TYB4_RABIT
ID TYB4_RABIT STANDARD; PRT; 43 AA.
AC P34032;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Thymosin beta-4 (T beta 4).
GN TMSB4.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE.
RX MEDLINE=83177074; PubMed=6838210;
```

RA Erickson-Viitanen S., Ruggieri S., Natalini P., Horecker B.L.;
RL "Distribution of thymosin beta 4 in vertebrate classes.";
RT Arch. Biochem. Biophys. 221:570-576(1983).
CC -1- FUNCTION: Plays an important role in the organization of the
CC cytoskeleton. Binds to and sequesters actin monomers (G actin) and
CC therefore inhibits actin polymerization (By similarity).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- TISSUE SPECIFICITY: ORIGINALLY FOUND IN THYMUS BUT IT IS WIDELY
CC DISTRIBUTED IN MANY TISSUES.
CC -1- SIMILARITY: BELONGS TO THE THYMOSIN BETA FAMILY.
DR HSSP: P21752: 1HJ0.
DR InterPro: IPR001152; Thymosin_b4.
DR Pfam: PF01290; Thymosin; 1.
DR ProDom: PD005116; Thymosin_b4; 1.
DR SMART: SM00152; THY; 1.
DR PROSITE: PS00500; THYMOSIN_B4; 1.
KW Actin-binding; Cytoskeleton; Acetylation.
FT MOD_RES 1 1 ACETYLATION.
SQ SEQUENCE 43 AA; 4905 MW; 440C76B45C6B03D0 CRC64;
Query Match 44.8%; Score 95.5; DB 1; Length 43;
Best Local Similarity 60.5%; Pred. No. 7.5e-05;
Matches 23; Conservative 3; Mismatches 11; Indels 1; Gaps 1;
QY 2 AHKLDLEETASLDKAKLKATEMQ-KNTLMTKETTEQEK 38
DB 1 ADKPDMAIEKFDKSLKKTETQEKNPPLPSKETIEQEK 38
RESULT 11
TYB4_XENLA
ID TYB4_XENLA STANDARD; PRT; 43 AA.
AC P18756;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Thymosin beta-4 (T beta 4) (Thymosin beta 4Xen).
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92231989; PubMed=1567461;
RA Yamamoto M., Shoda A., Minamino N., Matsuo H., Nishimatsu S., Ueno N.,
RA Murakami K.;
RT "Expression of thymosin beta 4 gene during Xenopus laevis
RT embryogenesis.";
RL Biochem. Biophys. Res. Commun. 184:93-99(1992).
RN [2]
RP SEQUENCE.
RC TISSUE=Oocyte;
RX MEDLINE=88132867; PubMed=3124756;
RA Hannappel E., Kalbacher H., Voelter W.;
RT "Thymosin beta 4Xen: a new thymosin beta 4-like peptide in oocytes of
RT Xenopus laevis.";
RL Arch. Biochem. Biophys. 260:546-551(1988).
CC -1- FUNCTION: Plays an important role in the organization of the
CC cytoskeleton. Binds to and sequesters actin monomers (G actin) and
CC therefore inhibits actin polymerization (By similarity).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- TISSUE SPECIFICITY: SPLEEN, KIDNEY, HEART, AND OOCYTES.
CC -1- SIMILARITY: BELONGS TO THE THYMOSIN BETA FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; D10692; BAA01534.1; -
DR PIR; S06344; S06344.
DR PIR; JQ1489; JQ1489.
DR HSSP: P21752: 1HJ0.
DR InterPro: IPR001152; Thymosin_b4.
DR Pfam: PF01290; Thymosin; 1.
DR ProDom: PD005116; Thymosin_b4; 1.
DR SMART: SM00152; THY; 1.
DR PROSITE: PS00500; THYMOSIN_B4; 1.
KW Actin-binding; Cytoskeleton; Acetylation.
FT INIT_MET 0 0
FT MOD_RES 1 1 ACETYLATION.
FT CONFLICT 40 41 TS -> ST (in ref. 1).
SQ SEQUENCE 43 AA; 4965 MW; 440C7CDEB4301D0 CRC64;
Query Match 44.8%; Score 95.5; DB 1; Length 43;
Best Local Similarity 63.9%; Pred. No. 7.5e-05;
Matches 23; Conservative 2; Mismatches 10; Indels 1; Gaps 1;
QY 4 KLDLEEIASLDKAKLKATEMQ-KNTLMTKETTEQEK 38
DB 3 KPDMAIEKFDKSLKKTETQEKNPPLPSKETIEQEK 38
RESULT 12
TYB_BRARE
ID TYB_BRARE STANDARD; PRT; 44 AA.
AC Q9W7M8;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Thymosin beta.
OS Brachydanio rerio (zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99169080; PubMed=10068630;
RA Roth L.W., Bormann P., Bonnet A., Reinhard E.;
RT "Beta-thymosin is required for axonal tract formation in developing
RT zebrafish brain.";
RL Development 126:1365-1374(1999).
CC -1- FUNCTION: Plays an important role in the organization of the
CC cytoskeleton. Binds to and sequesters actin monomers (G actin) and
CC therefore inhibits actin polymerization (By similarity).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: BELONGS TO THE THYMOSIN BETA FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; AF006831; AAD42010.1; -
DR InterPro: IPR001152; Thymosin_b4.
DR Pfam: PF01290; Thymosin; 1.
DR ProDom: PD005116; Thymosin_b4; 1.
DR SMART: SM00152; THY; 1.
DR PROSITE: PS00500; THYMOSIN_B4; 1.
KW Actin-binding; Cytoskeleton.
FT INIT_MET 0 0 BY SIMILARITY.
SQ SEQUENCE 44 AA; 5077 MW; 91F56A206956BC4A CRC64;
Query Match 44.8%; Score 95.5; DB 1; Length 44;
Best Local Similarity 56.1%; Pred. No. 7.7e-05;
Matches 23; Conservative 4; Mismatches 13; Indels 1; Gaps 1;
QY 2 AHKLDLEETASLDKAKLKATEMQ-KNTLMTKETTEQEK 41


```

DR PRODOM; PD005116; Thymosin_b4; 1.
DR SMART; SM00152; THY; 1.
DR PROSITE; PS00500; THYMOSIN_B4; 1.
KW Actin-binding; Cytoskeleton; Acetylation.
FT INIT_MET 0 0
FT PEPTIDE 1 4 HEMATOPOIETIC SYSTEM REGULATORY PEPTIDE.
FT MOD_RES 1 1 ACETYLATION.
FT CONFLICT 7 7 A -> V (IN REF. 5).
SQ SEQUENCE 43 AA; 4921 MW; 440C6102871103D0 CRC64;

Query Match
Best Local Similarity 43.4%; Score 92.5; DB 1; Length 43;
Matches 22; Conservative 3; Mismatches 10; Indels 1; Gaps 1;

QY 4 KLDLEETASLDKAKIKATEM-QNTLMTKETTEQEK 38
DB 3 KPDMAIEKFDKSKLTKTETQEKNTPLSKETIQEK 38

RESULT 14
TYBN_COTJA
ID TYBN_COTJA STANDARD; PRT; 44 AA.
AC Q9DET5;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DE Thymosin beta.
OS Coturnix coturnix japonica (Japanese quail)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianinae;
OC Coturnix.
OX NCBI_TaxID=93934;
RN [1]
RP SEQUENCE FROM N.A.
RA Dabthe V.E., Prols F., Brand-Saberi B.;
RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Plays an important role in the organization of the
CC cytoskeleton. Binds to and sequesters actin monomers (G actin) and
CC therefore inhibits actin polymerization (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: BELONGS TO THE THYMOSIN BETA FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AJ301650; CAC17010.1;
CC InterPro; IPR001152; Thymosin_b4.
CC Pfam; PF01290; Thymosin; 1.
CC SMART; SM00152; THY; 1.
CC PROSITE; PS00500; THYMOSIN_B4; 1.
CC Actin-binding; Cytoskeleton.
FT INIT_MET 0 0 BY SIMILARITY.
SQ SEQUENCE 44 AA; 5114 MW; 653D099D1D9AAC62 CRC64;

Query Match
Best Local Similarity 42.5%; Score 90.5; DB 1; Length 44;
Matches 21; Conservative 3; Mismatches 11; Indels 1; Gaps 1;

QY 4 KLDLEETASLDKAKIKATEM-QKNTLMTKETTEQEK 38
DB 3 KPDLESEKFDKSKLTKTETQEKNTPLSKETIQEK 38

RESULT 15
TYBN_HUMAN
ID TYBN_HUMAN STANDARD; PRT; 44 AA.
AC Q99406;
DT 15-DEC-1998 (Rel. 37, Created)

```

```

DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE NB Thymosin beta.
GN TMSNB.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=97191543; PubMed=9039501;
RA Yokoyama M., Nishi Y., Yoshii J., Okubo K., Matsubara K.;
RT Identification and cloning of neuroblastoma-specific and nerve
RT tissue-specific genes through compiled expression profiles.;
RL DNA Res. 3:311-320(1996).
RN [2]
RP SEQUENCE FROM N.A.
RA Bird C.;
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Plays an important role in the organization of the
CC cytoskeleton. Binds to and sequesters actin monomers (G actin) and
CC therefore inhibits actin polymerization (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- TISSUE SPECIFICITY: NEUROBLASTOMA-SPECIFIC.
CC -!- SIMILARITY: BELONGS TO THE THYMOSIN BETA FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; DB2345; BAA11556.1;
CC EMBL; AL035609; CAC18959.1;
CC HSSP; EC000183; AAH00183.1;
CC InterPro; IPR001152; Thymosin_b4.
CC Pfam; PF01290; Thymosin; 1.
CC SMART; SM00152; THY; 1.
CC PROSITE; PS00500; THYMOSIN_B4; 1.
CC Actin-binding; Cytoskeleton; Acetylation.
FT INIT_MET 0 0 BY SIMILARITY.
FT MOD_RES 1 1 ACETYLATION (BY SIMILARITY).
SQ SEQUENCE 44 AA; 5098 MW; B118286809C2656B CRC64;

Query Match
Best Local Similarity 40.6%; Score 86.5; DB 1; Length 44;
Matches 19; Conservative 6; Mismatches 10; Indels 1; Gaps 1;

QY 4 KLDLEETASLDKAKIKATEM-QKNTLMTKETTEQEK 38
DB 3 KPDLESEKFDKSKLTKTETQEKNTPLSKETIQEK 38

Search completed: May 30, 2003, 10:32:03
Job time : 23 secs

```

R;McCreary, V.; Kartha, S.; Bell, G.I.; Toback, F.G.

A:Accession: A60290
A:Molecule type: protein
A:Residues: 1-41 <LOW>
R:Hannappel, E.; Wartenberg, F.; Bustelo, X.R.
Arch. Biochem. Biophys. 273, 396-402, 1989
A:Title: Isolation and characterization of thymosin beta(9)(Met) from pork spleen
C:Keywords: acetylated amino end
A:Reference number: S05644; MUID:89372897; PMID:2774558
A:Accession: S05644
A:Molecule type: protein
A:Residues: 1-41 <HAN>
C:Superfamily: thymosin beta
C:Keywords: acetylated amino end
F:/Modified site: acetylated amino end (Ala) #status predicted

Query Match 51.9%; Score 110.5; DB 1; Length 41;
Best Local Similarity 71.1%; Pred. No. 1.6e-06;
Matches 27; Conservative 1; Mismatches 9; Indels 1; Gaps 1;

QY 2 AHKLDLEETASLDKAKLKATEMO-KNTLMTKETTEQEOK 38
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB 1 ADKPDGMEINFSFDKALKKTETQEKNTLPKYETIEQEOK 38

RESULT 5
S21282
thymosin beta-11 - rainbow trout
C:Species: Oncorhynchus mykiss (rainbow trout)
C:Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 31-Oct-1997
C:Accession: S21282
R:Vialouris, P.P.; Coles, B.; Tsitsiloni, O.; Schmid, B.; Howell, S.; Aitken, A.;
Biochem. J. 283, 385-389, 1992
A:Title: The complete sequences of trout (Salmo gairdneri) thymosin beta(11) and
A:Reference number: S21282; MUID:92246862; PMID:1575682
A:Accession: S21282
A>Status: preliminary
A:Molecule type: protein
A:Residues: 1-43 <VIA>
C:Superfamily: thymosin beta

Query Match 50.0%; Score 106.5; DB 2; Length 43;
Best Local Similarity 69.4%; Pred. No. 5e-06;
Matches 25; Conservative 2; Mismatches 8; Indels 1; Gaps 1;

QY 4 KLDLEETASLDKAKLKATEMO-KNTLMTKETTEQEOK 38
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB 5 KPNEEVASFDTKLKKTKTETQEKNPFPKYETIEQEOK 40

RESULT 6
A59005
thymosin beta - sea urchin (Arbacia punctulata)
C:Species: Arbacia punctulata (punctureate urchin)
C:Date: 01-Aug-1997 #sequence_revision 01-Aug-1997 #text_change 24-Nov-1999
C:Accession: A59005
R:Safer, D.; Chowrashi, P.K.
submitted to the Protein Sequence Database, July 1997
A:Reference number: A59005
A:Accession: A59005
A>Status: preliminary
A:Molecule type: protein
A:Residues: 1-40 <SAF>
C:Superfamily: thymosin beta
C:Keywords: acetylated amino end; actin binding
F:/Modified site: acetylated amino end (Ala) #status experimental

Query Match 49.1%; Score 104.5; DB 2; Length 40;
Best Local Similarity 60.5%; Pred. No. 7.9e-06;
Matches 23; Conservative 6; Mismatches 8; Indels 1; Gaps 1;

QY 2 AHKLDLEETASLDKAKLKATEMO-KNTLMTKETTEQEOK 38
DB 1 ADKPDPVSFTSDPKSLKLTQETQENKTLPTKDTIEQEOK 38


```
RESULT 7
JQ1489
thymosin beta-4 - African clawed frog
C:Species: Xenopus laevis (African clawed frog)
C>Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 16-Jun-2000
C:Accession: JQ1489; PO0330; S06344
R:Yanamoto, M.; Shoda, A.; Minamino, N.; Matsuo, H.; Nishimatsu, S.; Ueno, N.; Murakami,
Biochem. Biophys. Res. Commun. 184, 93-99, 1992
A:Title: Expression of thymosin beta-4 gene during Xenopus laevis embryogenesis.
A:Reference number: JQ1489; MUID:92231989; PMID:1567461
A:Accession: JQ1489
A:Molecule type: mRNA
A:Residues: 1-44 <YAM>
A:Cross-references: GB:D10692; NID:g222972; PIDN:BAA01534.1; PID:g222973
A:Accession: PO0330
A:Molecule type: protein
A:Residues: 2-15;20-44 <YA>
A:Experimental source: embryo
A:Note: the authors translated the codon for residue 6 as Lys
R:Hannappel, E.; Kalbacher, H.; Voelter, W.
Arch. Biochem. Biophys. 260, 546-551, 1988
A:Title: Thymosin-beta(Xen)(4): a new thymosin beta(4)-like peptide in oocytes of Xenopus
A:Reference number: S06344; MUID:88132867; PMID:3124756
A:Accession: S06344
A:Molecule type: protein
A:Residues: 2-40,'TS',43-44 <HAN>
C:Superfamily: thymosin beta
C:Keywords: actin binding; blocked amino end
F:2-44/Product: thymosin beta-4 #status experimental <MAT>
F:2/Modified site: blocked amino end (Ser) (in mature form) (probably acetylated) #status
experimental
Query Match 48.1%; Score 102.5; DB 1; Length 44;
Best Local Similarity 61.0%; Pred. No. 1.5e-05;
Matches 25; Conservative 3; Mismatches 12; Indels 1; Gaps 1;
QY 1 MAHKLDLEIASLDKAKLKATEMQ-KNTLMTKETTEQEK 40
Db 1 MSDKPDMAEIEKFDKAKLKKTETQEKNPPLSKETIEQEK 41
RESULT 8
B59005
thymosin beta - scallop (Argopecten irradians)
C:Species: Argopecten irradians
C>Date: 01-Aug-1997 #sequence_revision 01-Aug-1997 #text_change 24-Nov-1999
C:Accession: B59005
R:Safer, D.; Chowrashi, P.K.
Submitted to the Protein Sequence Database, July 1997
A:Reference number: A59005
A:Accession: B59005
A:Status: preliminary
A:Molecule type: protein
A:Residues: 1-40 <SAF>
A:Note: the order of the first four residues was assigned by homology to other beta-thym
C:Superfamily: thymosin beta
C:Keywords: acetylated amino end; actin binding
F:1/Modified site: acetylated amino end (Ser) #status experimental
Query Match 47.7%; Score 101.5; DB 2; Length 40;
Best Local Similarity 61.1%; Pred. No. 1.8e-05;
Matches 22; Conservative 6; Mismatches 7; Indels 1; Gaps 1;
QY 4 KLDLEIASLDKAKLKATEP-MQKNTLMTKETTEQEK 38
Db 3 KPDVSEVANFDKSKLKKTETAEKNLPLTKETIQEK 38
RESULT 9
S22426
thymosin beta-12 - rainbow trout
C:Species: Oncorhynchus mykiss (rainbow trout)
C>Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 31-Oct-1997
```

```
C:Accession: S22426
R:Yialouris, P.P.; Coles, B.; Tsitsiloni, O.; Schmid, B.; Howell, S.; Aitken, A.;
Biochem. J. 283, 385-389, 1992
A:Title: The complete sequences of trout (Salmo gairdneri) thymosin beta(11) and thymosin
A:Reference number: S21282; MUID:92246862; PMID:1575682
A:Accession: S22426
A:Status: preliminary
A:Molecule type: protein
A:Residues: 1-44 <VIA>
C:Superfamily: thymosin beta
Query Match 46.7%; Score 99.5; DB 2; Length 44;
Best Local Similarity 63.9%; Pred. No. 3.4e-05;
Matches 23; Conservative 3; Mismatches 9; Indels 1; Gaps 1;
QY 4 KLDLEIASLDKAKLKATEMQ-KNTLMTKETTEQEK 38
Db 5 KPDLAEVSNFDKTKLKKTETQEKNPPLTKETIEQEK 40
RESULT 10
A38682
thymosin beta-4 [validated] - human
N:Alternate names: actin-sequestering peptide Fx
C:Species: Homo sapiens (man)
C>Date: 06-Mar-1992 #sequence_revision 06-Sep-1996 #text_change 08-Dec-2000
C:Accession: I56000; A38682; I37581
R:Gondo, H.; Kudo, J.; White, J.W.; Barr, C.L.; Selvanayagam, P.; Saunders, G.F.
J. Immunol. 139, 3840-3848, 1987
A:Title: Differential expression of the human thymosin beta-4 gene in lymphocytes,
A:Reference number: I56000; MUID:88060494; PMID:3500230
A:Accession: I56000
A:Status: translated from GB/EMBL/DDBJ
A:Molecule type: mRNA
A:Residues: 1-44 <RES>
A:Cross-references: GB:M17733; NID:g339688; PIDN:AAA36745.1; PID:g339689
R:Safer, D.; Elzinga, M.; Nachmias, V.T.
J. Biol. Chem. 266, 4029-4032, 1991
A:Title: Thymosin beta-4 and Fx, an actin-sequestering peptide, are indistinguishable
A:Reference number: A38682; MUID:91154186; PMID:1999398
A:Accession: A38682
A:Molecule type: protein
A:Residues: 2-44 <SAF>
R:Friedman, R.L.; Manly, S.P.; McMahon, M.; Kerr, I.M.; Stark, G.R.
Cell 38, 745-755, 1984
A:Title: Transcriptional and posttranscriptional regulation of interferon-induced
A:Reference number: I37581; MUID:85024867; PMID:6548414
A:Accession: I37581
A:Status: translated from GB/EMBL/DDBJ
A:Molecule type: mRNA
A:Residues: 8-44 <RE2>
A:Cross-references: EMBL:X02493; NID:g32703; PIDN:CAA26323.1; PID:g825683
C:Comment: This protein was once considered a thymic hormone. While the possibility of
the binding and sequestering of actin monomers.
C:Genetics:
A:Gene: GDB:TMSB4
A:Cross-references: GDB:378616; OMIM:188395
A:Map position: 12q22-12q22
C:Superfamily: thymosin beta
C:Keywords: actin binding
F:2-44/Product: thymosin beta-4 #status experimental <MAT>
Query Match 45.8%; Score 97.5; DB 1; Length 44;
Best Local Similarity 59.0%; Pred. No. 5.8e-05;
Matches 23; Conservative 4; Mismatches 11; Indels 1; Gaps 1;
QY 1 MAHKLDLEIASLDKAKLKATEMQ-KNTLMTKETTEQEK 38
Db 1 MSDKPDMAEIEKFDKSKLKKTETQEKNPPLSKETIEQEK 39
RESULT 11
A37217
```

thymosin beta-4 - mouse
C:Species: Mus musculus (house mouse)
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C:Accession: A37217; S08074; S12884
R:Low, T.L.K.; Engler, P.; Storb, U.
J. Immunol. 144, 4857-4862, 1990.
A:Title: Differential splicing of thymosin beta-4 mRNA.
A:Reference number: A37217; MUID:90278100; PMID:2351831
A:Accession: A37217
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-50 <RUD>
A:Cross-references: GB:X16053; GB:M54991; NID:954793; PIDN:CAA34187.1; PID:g54794; GB:M54991
R:Rudin, C.; Engler, P.; Provost, S.; Storb, U.
submitted to the EMBL Data Library, August 1989
A:Reference number: S08074
A:Accession: S08074
A:Molecule type: mRNA
A:Residues: 1-50 <RUD>
A:Cross-references: EMBL:X16053; NID:954793; PIDN:CAA34187.1; PID:g54794
R:Low, T.L.K.; Pan, T.L.; Lin, Y.S.
FEBS Lett. 273, 1-5, 1990
A:Title: Depression of prothymosin-alpha production in murine thymus correlates with status
A:Reference number: S12883; MUID:91032153; PMID:2226839
A:Accession: S12884
A:Status: preliminary
A:Molecule type: protein
A:Residues: 8-50 <LOW>
C:Superfamily: thymosin beta-4
C:Keywords: actin binding; alternative splicing
F:1-50/Product: thymosin beta-4 precursor, long splice form #status predicted <LFO>
F:7-50/Product: thymosin beta-4 precursor, short splice form #status predicted <LFO>
F:8-50/Product: thymosin beta-4 #status experimental <MAT>
Query Match
Best Local Similarity 45.8%; Score 97.5; DB 1; Length 50;
Matches 23; Conservative 4; Mismatches 11; Indels 1; Gaps 1;
QY 1 MAHKLDLEETASLDKAKLKATEMQ-KNTLMTKETTEQEK 38
Db 7 MSKDPDMAETEKEDKSLKATETQEKNPPLPSKETIEQEK 45
RESULT 12
I52084
thymosin beta-4 precursor - rat (fragment)
C:Species: Rattus norvegicus (Norway rat)
C:Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 16-Jul-1999
C:Accession: I52084
R:Goodall, G.J.; Richardson, M.; Furuichi, Y.; Wodnar-Fillipowicz, A.; Horecker, B.L.
Arch. Biochem. Biophys. 236, 445-447, 1985
A:Title: Sequence of a cloned 523-bp cDNA for thymosin beta-4.
A:Reference number: I52084; MUID:85095883; PMID:3838131
A:Accession: I52084
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: mRNA
A:Residues: 1-56 <RES>
A:Cross-references: GB:M26759; NID:g207317; PIDN:AAA42246.1; PID:g207318
C:Superfamily: thymosin beta-4
Query Match
Best Local Similarity 45.8%; Score 97.5; DB 2; Length 56;
Matches 23; Conservative 4; Mismatches 11; Indels 1; Gaps 1;
QY 1 MAHKLDLEETASLDKAKLKATEMQ-KNTLMTKETTEQEK 38
Db 13 MSKDPDMAETEKEDKSLKATETQEKNPPLPSKETIEQEK 51
RESULT 13
S21694
thymosin beta-12 - Japanese seaperch
C:Species: Lateolabrax japonicus (Japanese seaperch)

thymosin beta-4 - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 15-Nov-1984 #sequence_revision 15-Nov-1984 #text_change 18-Jun-1999
C:Accession: A01522; A36123
R:Wodnar-Fillipowicz, A.; Gubler, U.; Furuichi, Y.; Richardson, M.; Nowoswiat, E.F.;
Proc. Natl. Acad. Sci. U.S.A. 81, 2295-2297, 1984
A:Title: Cloning and sequence analysis of cDNA for rat spleen thymosin beta-4.
A:Reference number: A01522; MUID:84193997; PMID:6201851
A:Accession: A01522
A:Molecule type: mRNA
A:Residues: 1-44 <WOD>
A:Cross-references: GB:K01334; NID:g207315; PIDN:AAA42245.1; PID:g207316
R:Atkinson, M.J.; Freeman, M.W.; Kronenberg, H.M.
Mol. Endocrinol. 4, 69-74, 1990
A:Title: Thymosin beta-4 is expressed in ROS 17/2.8 osteosarcoma cells in a regulated
A:Reference number: A36123; MUID:90220652; PMID:2325669
A:Accession: A36123
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-7, 'A', '9-44 <ATR>
A:Cross-references: GB:M34043; NID:g206707; PIDN:AAA42062.1; PID:g206708
C:Comment: This protein does not have the typical amino-terminal hydrophobic region u
C:Superfamily: thymosin beta-4
C:Keywords: actin binding
Query Match
Best Local Similarity 45.3%; Score 96.5; DB 1; Length 44;
Matches 23; Conservative 4; Mismatches 11; Indels 1; Gaps 1;
QY 1 MAHKLDLEETASLDKAKLKATEMQ-KNTLMTKETTEQEK 38
Db 1 MSKDPDMAETEKEDKSLKATETQEKNPPLPSKETIEQEK 39
RESULT 15
TNB084
thymosin beta-4 - bovine
N:Alternate names: hematopoietic stem cell proliferation inhibitor
C:Species: Bos primigenius taurus (cattle)
C:Date: 30-Apr-1981 #sequence_revision 30-Apr-1981 #text_change 24-Feb-1995
C:Accession: A01521; A31451; S13193
R:Low, T.L.K.; Hu, S.K.; Goldstein, A.L.
Proc. Natl. Acad. Sci. U.S.A. 78, 1162-1166, 1981
A:Title: Complete amino acid sequence of bovine thymosin beta-4: a thymic hormone tha
A:Reference number: A01521; MUID:81199361; PMID:6940133
A:Accession: A01521
A:Molecule type: protein
A:Residues: 1-43 <LOW>
R:Lenfant, M.; Wdziecjak-Bakala, J.; Guittet, E.; Prome, J.C.; Sotty, D.; Frindel, E.

Query Match 43.4%; Score 92.5; DB 1; Length 43;
Best Local Similarity 61.1%; Pred. No. 0.00022;
Matches 22; Conservative 3; Mismatches 10; Indels 1; Gaps 1;

Search completed: May 30, 2003, 10:35:32
Job time : 42 secs

THIS PAGE BLANK (J.3870)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 30, 2003, 10:31:31 ; Search time 70 Seconds
(without alignments)
81.854 Million cell updates/sec

Title: US-09-915-178-2

Perfect score: 213
Sequence: 1 MAHKLDLEETASLDKAKLKA.....OKNTLMTKETTEQKWSBIS 43

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_101002.*
1: /SID22/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SID22/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SID22/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SID22/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SID22/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SID22/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SID22/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SID22/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SID22/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SID22/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SID22/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SID22/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SID22/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SID22/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SID22/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SID22/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SID22/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SID22/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SID22/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SID22/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SID22/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SID22/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SID22/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	213	100.0	43	ABB05034	Human NOV2a protei
2	213	100.0	43	ABB05035	Human NOV2b protei
3	213	100.0	43	ABB05036	Human NOV2c protei
4	130.5	61.3	43	AA196932	Thymosin beta 10.
5	130.5	61.3	43	AA196932	Thymosin beta 4 pe
6	128	60.1	58	AAU00689	Thymosin-beta-10-1
7	128	60.1	58	AAU00689	Thymosin-beta-10-1
8	127.5	59.9	43	AA196935	N-Terminus of thym
9	116.5	54.7	41	AA196935	Thymosin beta 13.
10	116.5	54.7	41	AA196935	Thymosin beta 4 pe

11	112.5	52.8	39	13	AA196935	N-Terminus of thym
12	112.5	52.8	41	13	AA196935	N-Terminus of thym
13	112.5	52.8	41	17	AA196935	Thymosin beta 9.
14	112.5	52.8	41	21	AA196935	Thymosin beta 4 pe
15	110.5	51.9	41	17	AA196935	Thymosin beta 9-NE
16	110.5	51.9	41	21	AA196935	Thymosin beta 4 pe
17	106.5	50.9	41	13	AA196935	N-Terminus of thym
18	106.5	50.9	41	17	AA196935	N-Terminus of thym
19	106.5	50.9	41	17	AA196935	Thymosin beta 11.
20	104.5	49.1	40	17	AA196935	Thymosin beta 4 pe
21	104.5	49.1	40	17	AA196935	Thymosin beta 14.
22	103.5	48.6	44	23	ABP60994	Novel human protei
23	103.5	48.6	74	22	ABP60994	Novel human protei
24	99.5	46.7	40	21	AA196935	Thymosin beta 4 pe
25	99.5	46.7	42	17	AA196935	Thymosin beta 12.
26	99.5	46.7	42	21	AA196935	Thymosin beta 4 pe
27	98.5	46.2	58	21	AA196935	Human secreted pro
28	97.5	45.8	43	21	AA196935	Thymosin beta 4 pe
29	97.5	45.8	44	19	AA196935	Thymosin beta 4 pe
30	97.5	45.8	56	21	AA196935	Human colon cancer
31	97.5	45.8	57	21	AA196935	Human cytoskeleton
32	97.5	45.8	68	9	AA196935	Protein produced i
33	96.5	45.3	43	21	AA196935	Thymosin beta 4 pe
34	95.5	44.8	43	17	AA196935	Thymosin beta 4 pe
35	95.5	44.8	43	17	AA196935	Thymosin beta 4-Xe
36	95.5	44.8	45	18	AA196935	Human thymosin bet
37	95.5	44.8	45	19	AA196935	Rat thymosin-beta1
38	95.5	44.8	45	19	AA196935	Human thymosin bet
39	95.5	44.8	45	19	AA196935	Human thymosin bet
40	95.5	44.8	45	21	AA196935	Human thymosin bet
41	95.5	44.8	45	23	AA196935	Human thymosin bet
42	93.5	43.9	40	21	AA196935	Thymosin beta 4 pe
43	92.5	43.4	43	2	AA196935	Thymosin beta-4.
44	92.5	43.4	43	17	AA196935	Thymosin beta 4 pe
45	92.5	43.4	43	21	AA196935	Thymosin beta 4 pe

ALIGNMENTS

RESULT 1		ABB05034		standard; Protein; 43 AA.	
ID	ABB05034	standard; Protein; 43 AA.			
XX	XX	AC			
XX	XX	ABB05034;			
DT	25-MAR-2002	(first entry)			
XX	XX	Human NOV2a protein SEQ ID NO:4.			
DE	DE	Human; NOVX; cytostatic; antidiabetic; anorectic; antibacterial;			
KW	KW	fungicide; virucide; protozoacide; analgesic; antiparkinsonian;			
KW	KW	antithrombotic; hypotensive; osteoprotective; antidiabetic;			
KW	KW	neuroprotective; cardiatic; antiallergic; antidiabetic; nootropic;			
KW	KW	anticonvulsant; neuroleptic; gene therapy; vaccine; immune disease;			
KW	KW	developmental disease; taste and scent detectability disorder; infection;			
KW	KW	Burkitt's lymphoma; signal transduction pathway disorder; pain; cancer;			
KW	KW	retinal disease; feeding disorder; asthma; Parkinson's disease; ulcer;			
KW	KW	noninsulin-dependent diabetes mellitus; acute heart failure; hypotension;			
KW	KW	hypertension; urinary retention; osteoporosis; Crohn's disease; allergy;			
KW	KW	multiple sclerosis; angina pectoris; myocardial infarction; delirium;			
KW	KW	benign prostatic hypertrophy; manic depression; dementia; dyskinesia;			
KW	KW	severe mental retardation; Huntington's disease; chromosome 9;			
XX	XX	Gilles de la Tourette syndrome.			
OS	OS	Homo sapiens.			
XX	XX	WO200190155-A2.			
XX	XX	29-NOV-2001.			
XX	XX	24-MAY-2001; 2001WO-US17073.			
XX	XX				

PR 24-MAY-2000; 2000US-206679P.
PR 24-MAY-2000; 2000US-206688P.
PR 24-MAY-2000; 2000US-206829P.
PR 30-MAY-2000; 2000US-207748P.
PR 30-MAY-2000; 2000US-207798P.
PR 31-MAY-2000; 2000US-208263P.
PR 02-JUN-2000; 2000US-208831P.
PR 05-JUN-2000; 2000US-209451P.
PR 07-JUN-2000; 2000US-210060P.
PR 20-JUL-2000; 2000US-219507P.
PR 26-JUL-2000; 2000US-221337P.
PR 31-JUL-2000; 2000US-221927P.
PR 19-JAN-2001; 2001US-263135P.
PR 24-JAN-2001; 2001US-263688P.
PR 24-JAN-2001; 2001US-263694P.
PR 23-MAY-2001; 2001US-0863776.
XX (CURA-) CURAGEN CORP.
XX Spytek KA, Majumder K, Tchernev VT, Mishra V, Padigaru M;
XX Spaderna SK, Shenoy S, Rastelli L, Li L, Taupier RJ, Gangolli E;
XX WPI: 2002-106174/14.
XX N-PSDB; ABA92649.
XX Novel polypeptide, useful for treating pain, cancer, urinary retention,
PT osteoporosis, Crohn's disease, diabetes, acute heart failure, dementia,
PT asthma, ulcer, allergy and Huntington's disease, comprises isolated
PT polypeptide NOVX -
XX Claim 1; page 20; 266pp; English.
XX The present invention describes human NOVX proteins (NOV1 to NOV9). NOVX
CC proteins (I) have cytostatic, antidiabetic, anorectic, antibacterial,
CC fungicide, virucide, protozoacide, analgesic, antiparkinsonian,
CC antiasthmatic, hypotensive, osteopathic, antiinflammatory, antiulcer,
CC neuroprotective, cardiant, antiallergic, antidepressant, nootropic,
CC anticonvulsant and neuroleptic activities. (I) and (II) can be
CC used in gene therapy and vaccine production. (I) and (II) can be
CC used for treating or preventing a NOVX-associated disorder such as
CC cardiomyopathy, atherosclerosis and diabetes in a human, where the
CC disorder is related to cell signal processing and metabolic path way
CC modulation, in a subject, preferably human. (I) and (II) can be used for
CC diagnosing, preventing or treating developmental diseases, immune
CC signal transduction pathway disorders, retinal diseases including those
CC involving photoreception, cell growth rate disorders, feeding disorders,
CC noninsulin-dependent diabetes mellitus, infections, pain, cancer, asthma,
CC Parkinson's disease, acute heart failure, hypotension, hypertension,
CC urinary retention, osteoporosis, Crohn's disease, multiple sclerosis,
CC Albright Hereditary Osteodystrophy, angina pectoris, ulcers, myocardial
CC infarction, allergies, benign prostatic hypertrophy, manic depression,
CC delirium, dementia, severe metal retardation and dyskinesias, such as
CC Huntington's disease or Gilles de la Tourette syndrome. The present
CC sequence represents the human NOV2a protein from the present invention.
XX Human NOV2a is isolated to chromosome 9.
XX Sequence 43 AA:
Query Match 100.0%; Score 213; DB 23; Length 43;
Best Local Similarity 100.0%; Pred. No. 2.7e-19;
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MAHKLDLEETASLDKAKLKATEMKQNTLMTKETTTQEKWSEIS 43
DB 1 MAHKLDLEETASLDKAKLKATEMKQNTLMTKETTTQEKWSEIS 43
RESULT 2
ABB05035
ID ABB05035 standard; Protein; 43 AA.
XX
AC ABB05035;

25-MAR-2002 (first entry)
Human NOV2b protein SEQ ID NO:6.
Human; NOVX; cytostatic; antidiabetic; anorectic; antibacterial;
fungicide; virucide; protozoacide; analgesic; antiparkinsonian;
antiasthmatic; hypotensive; osteopathic; antiinflammatory; antiulcer;
neuroprotective; cardiant; antiallergic; antidepressant; nootropic;
anticonvulsant; neuroleptic; gene therapy; vaccine; immune disease;
developmental disease; taste and scent detectability disorder; infection;
Burkitt's lymphoma; signal transduction pathway disorder; pain; cancer;
retinal disease; feeding disorder; asthma; Parkinson's disease; ulcer;
noninsulin-dependent diabetes mellitus; acute heart failure; hypotension;
hypertension; urinary retention; osteoporosis; Crohn's disease; allergy;
multiple sclerosis; angina pectoris; myocardial infarction; delirium;
benign prostatic hypertrophy; manic depression; dementia; dyskinesia;
severe metal retardation; Huntington's disease;
Gilles de la Tourette syndrome.
Homo sapiens.
WO200190155-A2.
29-NOV-2001.
24-MAY-2001; 2001WO-US17073.
24-MAY-2000; 2000US-206679P.
24-MAY-2000; 2000US-206688P.
24-MAY-2000; 2000US-206829P.
30-MAY-2000; 2000US-207748P.
30-MAY-2000; 2000US-207798P.
31-MAY-2000; 2000US-208263P.
02-JUN-2000; 2000US-208831P.
05-JUN-2000; 2000US-209451P.
07-JUN-2000; 2000US-210060P.
20-JUL-2000; 2000US-219507P.
26-JUL-2000; 2000US-221337P.
31-JUL-2000; 2000US-221927P.
19-JAN-2001; 2001US-263135P.
24-JAN-2001; 2001US-263688P.
24-JAN-2001; 2001US-263694P.
23-MAY-2001; 2001US-0863776.
(CURA-) CURAGEN CORP.
Spytek KA, Majumder K, Tchernev VT, Mishra V, Padigaru M;
Spaderna SK, Shenoy S, Rastelli L, Li L, Taupier RJ, Gangolli E;
WPI: 2002-106174/14.
N-PSDB; ABA92650.
Novel polypeptide, useful for treating pain, cancer, urinary retention,
osteoporosis, Crohn's disease, diabetes, acute heart failure, dementia,
asthma, ulcer, allergy and Huntington's disease, comprises isolated
polypeptide NOVX -
Claim 1; page 22; 266pp; English.
The present invention describes human NOVX proteins (NOV1 to NOV9). NOVX
proteins (I) have cytostatic, antidiabetic, anorectic, antibacterial,
fungicide, virucide, protozoacide, analgesic, antiparkinsonian,
antiasthmatic, hypotensive, osteopathic, antiinflammatory, antiulcer,
neuroprotective, cardiant, antiallergic, antidepressant, nootropic,
anticonvulsant and neuroleptic activities. (I) and (II) can be
used in gene therapy and vaccine production. (I) and (II) can be
used for treating or preventing a NOVX-associated disorder such as
cardiomyopathy, atherosclerosis and diabetes in a human, where the
disorder is related to cell signal processing and metabolic path way
modulation, in a subject, preferably human. (I) and (II) can be used for
diagnosing, preventing or treating developmental diseases, immune
signal transduction pathway disorders, retinal diseases including those
involving photoreception, cell growth rate disorders, feeding disorders,
noninsulin-dependent diabetes mellitus, infections, pain, cancer, asthma,
Parkinson's disease, acute heart failure, hypotension, hypertension,
urinary retention, osteoporosis, Crohn's disease, multiple sclerosis,
Albright Hereditary Osteodystrophy, angina pectoris, ulcers, myocardial
infarction, allergies, benign prostatic hypertrophy, manic depression,
delirium, dementia, severe metal retardation and dyskinesias, such as
Huntington's disease or Gilles de la Tourette syndrome. The present
sequence represents the human NOV2a protein from the present invention.
Human NOV2a is isolated to chromosome 9.

CC signal transduction pathway disorders, retinal diseases including those
 CC involving photoreception, cell growth rate disorders, feeding disorders,
 CC noninsulin-dependent diabetes mellitus, infections, pain, cancer, asthma,
 CC Parkinson's disease, acute heart failure, hypotension, hypertension,
 CC urinary retention, osteoporosis, Crohn's disease, multiple sclerosis,
 CC Albright Hereditary Osteodystrophy, angina pectoris, ulcers, myocardial
 CC infarction, allergies, benign prostatic hypertrophy, manic depression,
 CC delirium, dementia, severe metal retardation and dyskinesias, such as
 CC Huntington's disease or Gilles de la Tourette syndrome. The present
 CC sequence represents the human NOV2b protein from the present invention.

XX Sequence 43 AA;

Query Match 100.0%; Score 213; DB 23; Length 43;

Best Local Similarity 100.0%; Pred. No. 2.7e-19;

Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAHKLDLEETASLDKAKLKATEMOKNTLMTKETTEQKWEIS 43

DB 1 MAHKLDLEETASLDKAKLKATEMOKNTLMTKETTEQKWEIS 43

RESULT 3

ABB05036

ID ABB05036 standard; Protein; 43 AA.

XX AC

ABB05036;

DT 25-MAR-2002 (first entry)

XX Human NOV2c protein SEQ ID NO:8.

DE Human; NOVX; cytostatic; antidiabetic; anorectic; antibacterial;
 KW fungicide; virucide; protozoacide; analgesic; antiparkinsonian;
 KW antilasthmatic; hypotensive; osteopathic; antiinflammatory; antiulcer;
 KW neuroprotective; cardiant; antiallergic; antidepressant; nootropic;
 KW anticonvulsant; neuroleptic; gene therapy; vaccine; immune disease;
 KW developmental disease; taste and scent detectability disorder; infection;
 KW Burkitt's lymphoma; signal transduction pathway disorder; pain; cancer;
 KW retinal disease; feeding disorder; asthma; Parkinson's disease; ulcer;
 KW noninsulin-dependent diabetes mellitus; acute heart failure; hypotension;
 KW hypertension; urinary retention; osteoporosis; Crohn's disease; allergy;
 KW multiple sclerosis; angina pectoris; myocardial infarction; delirium;
 KW benign prostatic hypertrophy; manic depression; dementia; dyskinesia;
 KW severe metal retardation; Huntington's disease;
 KW Gilles de la Tourette syndrome.

XX Homo sapiens.

XX WO200190155-A2.

XX 29-NOV-2001.

XX 24-MAY-2001; 2001WO-US17073.

XX 24-MAY-2000; 2000US-206679P.

XX 24-MAY-2000; 2000US-206688P.

XX 24-MAY-2000; 2000US-206829P.

XX 30-MAY-2000; 2000US-207748P.

XX 30-MAY-2000; 2000US-207798P.

XX 31-MAY-2000; 2000US-208263P.

XX 02-JUN-2000; 2000US-208831P.

XX 05-JUN-2000; 2000US-209451P.

XX 07-JUN-2000; 2000US-210060P.

XX 20-JUL-2000; 2000US-219507P.

XX 26-JUL-2000; 2000US-221337P.

XX 31-JUL-2000; 2000US-221927P.

XX 19-JAN-2001; 2001US-263135P.

XX 24-JAN-2001; 2001US-263688P.

XX 24-JAN-2001; 2001US-263694P.

XX 23-MAY-2001; 2001US-0863776.

XX (CURA-) CURAGEN CORP.

XX

PI

XX

DR

DR

XX

PT

PT

PT

XX

PS

XX

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

Spyttek KA, Majumder K, Tchernev VT, Mishra V, Padigaru M;
 Spaderna SK, Shenoy S, Rastelli L, Li L, Taupier RJ, Gangolli E;
 WPI: 2002-106174/14.
 N-PSDB; ABA92651.

Novel polypeptide, useful for treating pain, cancer, urinary retention,
 osteoporosis, Crohn's disease, diabetes, acute heart failure, dementia,
 asthma, ulcer, allergy and Huntington's disease, comprises isolated
 polypeptide NOVX -

Claim 1; page 23; 266pp; English.

The present invention describes human NOVX proteins (NOV1 to NOV9). NOVX
 proteins (I) have cytostatic, antidiabetic, anorectic, antibacterial,
 fungicide, virucide, protozoacide, analgesic, antiparkinsonian,
 antiasthmatic, hypotensive, osteopathic, antiinflammatory, antiulcer,
 neuroprotective, cardiant, antiallergic, antidepressant, nootropic,
 anticonvulsant and neuroleptic activities. (I) and polynucleotides (II)
 can be used in gene therapy and vaccine production. (I) and (II) can be
 used for treating or preventing a NOVX-associated disorder such as
 cardiomyopathy, atherosclerosis and diabetes in a human, where the
 disorder is related to cell signal processing and diabetes in a human, where the
 modulation, in a subject, preferably human. (I) and (II) can be used for
 diagnosing, preventing or treating developmental diseases, immune
 diseases, taste and scent detectability disorder, Burkitt's lymphoma,
 signal transduction pathway disorders, retinal diseases including those
 involving photoreception, cell growth rate disorders, feeding disorders,
 noninsulin-dependent diabetes mellitus, infections, pain, cancer, asthma,
 Parkinson's disease, acute heart failure, hypotension, hypertension,
 urinary retention, osteoporosis, Crohn's disease, multiple sclerosis,
 Albright Hereditary Osteodystrophy, angina pectoris, ulcers, myocardial
 infarction, allergies, benign prostatic hypertrophy, manic depression,
 delirium, dementia, severe metal retardation and dyskinesias, such as
 Huntington's disease or Gilles de la Tourette syndrome. The present
 sequence represents the human NOV2c protein from the present invention.

Sequence 43 AA;

Query Match 100.0%; Score 213; DB 23; Length 43;

Best Local Similarity 100.0%; Pred. No. 2.7e-19;

Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAHKLDLEETASLDKAKLKATEMOKNTLMTKETTEQKWEIS 43

DB 1 MAHKLDLEETASLDKAKLKATEMOKNTLMTKETTEQKWEIS 43

RESULT 4

AAR96932

ID AAR96932 standard; protein; 43 AA.

XX AC

AAR96932;

XX 17-OCT-1996 (first entry)

XX Thymosin beta 10.

XX Thymosin beta 4; OAD; obstructive airway disease; sputum;

XX viscoelasticity; respiratory disorder; respiratory distress syndrome;

XX chronic bronchitis; asthma; emphysema; cystic fibrosis; actin;

XX polymerisation.

XX Synthetic.

XX WO9611016-A1.

XX 18-APR-1996.

XX 06-OCT-1995; 95WO-US12543.

XX 07-JUN-1995; 95US-0483572.

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

PR 06-OCT-1994; 94US-0318994.
PA (ALPH-) ALPHA 1 BIOMEDICALS INC.
XX (UTWA-) UNIV WASHINGTON MEDICAL CENT GEORGE.
PI Bauman C, Berman ML, Crockford DR, Goldstein AL;
PI Kater A, Rubin BK;
XX WPI; 1996-209656/21.
XX Reduction of viscoelasticity of sputum of obstructive airway disease
XX - for use in treating, e.g. cystic fibrosis.
XX Disclosure; Figure 1B; 45pp; English.
XX Thymosin beta 4 may be used in a composition alongside DNase1 to
XX reduce the viscoelasticity of sputum causing obstructive airway
XX disease (OAD). Such disease includes disorders such as acute and
XX chronic respiratory distress syndrome, chronic bronchitis, asthma,
XX emphysema and cystic fibrosis associated with excess actin
XX polymerisation (polymerisation of G-actin into F-actin).
XX Sequence 43 AA;
SQ

Query Match 61.3%; Score 130.5; DB 17; Length 43;
Best Local Similarity 74.4%; Pred. No. 4e-09;
Matches 32; Conservative 1; Mismatches 9; Indels 1; Gaps 1;
OY 2 AHKLDLEFIASLDKAKLKATEMO-KNTLMKTETEOKWSEIS 43
DB 1 ADKPDGMEIASFDKAKLKKTETQKNTLPTKETIEOKRSEIS 43
RESULT 5
ID AAY80267
XX AAY80267 standard; peptide: 43 AA.
AC AAY80267;
DT 26-MAY-2000 (first entry)
XX Thymosin beta 4 peptide isoform Tbeta10.
DE Thymosin beta 4
KW Thymosin beta 4; Tbeta4; tissue repair; angiogenesis; wound repair;
KW cell migration; vulnary; antiatherosclerotic; cytostatic; diagnosis;
KW healing disorder; fibrotic disease; ischaemia; atherosclerosis;
KW cell proliferative disorder.
XX Unidentified.
OS
XX
XX WO200006190-A1.
XX
XX 10-FEB-2000.
XX
XX 29-JUL-1999; 99WO-US17282.
XX
XX 30-JUL-1998; 98US-0094690.
XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX Kleinman HK, Goldstein A, Malinda KM, Sosne G;
XX WPI; 2000-195195/17.
XX
XX A novel method to promote wound healing comprises administering
XX Thymosin beta 4, useful for promoting tissue repair, angiogenesis and
XX cell migration
XX
XX Claim 10; Fig 11; 61pp; English.
XX
XX A method has been developed for promoting wound healing in a subject
XX comprising administering a composition containing a wound healing
XX polypeptide comprising the amino acid sequence Leu-Lys-Thr-Glu-Thr

CC and conservative variants, which have wound healing activity. The wound
CC healing polypeptide is preferably thymosin beta 4 (Tbeta4) or an isoform
CC of Tbeta4. The polypeptide has vulnary, antiatherosclerotic and
CC cytostatic activities. It is useful for promoting wound healing, in vivo
CC in a subject, especially humans, or ex vivo in tissues (e.g. urogenital,
CC gastro-intestinal, cardiovascular, muscle, connective and neural,
CC to inhibit/ameliorate wound healing. The level of Tbeta4 can be used
CC can be used to diagnose a pathological state characterised by a wound
CC healing disorder, such as fibrotic disease, ischaemia, atherosclerosis
CC and cell proliferative disorders. Tbeta4 or its isoforms can also be
CC used to promote epithelial cell (e.g. skin cells, especially
CC keratinocytes, corneal cells) migration and angiogenesis. The polypeptide
CC is able to promote wound repair without any side effects, unlike prior
CC art growth and angiogenic factors used for such treatment. The present
CC sequence represents a thymosin beta 4 peptide isoform from the present
XX invention.
XX
XX Sequence 43 AA;
SQ

Query Match 61.3%; Score 130.5; DB 21; Length 43;
Best Local Similarity 74.4%; Pred. No. 4e-09;
Matches 32; Conservative 1; Mismatches 9; Indels 1; Gaps 1;
OY 2 AHKLDLEFIASLDKAKLKATEMO-KNTLMKTETEOKWSEIS 43
DB 1 ADKPDGMEIASFDKAKLKKTETQKNTLPTKETIEOKRSEIS 43
RESULT 6
ID AAU00689
XX AAU00689 standard; Protein: 58 AA.
AC AAU00689;
XX
XX 07-SEP-2001 (first entry)
XX Thymosin-beta-10-like protein.
DE Thymosin-beta-10-like protein.
XX
XX Thymosin-beta-10-like protein; ephrin type-A receptor 8-like protein;
KW proteoglycan-like protein; fibromodulin; fibronectin; thymic immune cell;
KW spermatogenesis; male infertility; neoplasia; red blood cell; platelet;
KW small cell lung cancer; GPI-anchored ephrin-A ligand; prostate cancer;
KW neurological disorder; cardiac disorder; vascular disorder; orthopaedic;
KW inflammatory disease; rheumatoid arthritis; connective tissue;
XX congenital muscular dystrophy; chemotherapy; immunotherapy.
XX
XX Homo sapiens.
OS
XX
XX WO200129217-A2.
XX
XX 26-APR-2001.
XX
XX 13-OCT-2000; 2000WO-US28474.
XX
XX 15-OCT-1999; 99US-0159805.
XX 18-OCT-1999; 99US-0159992.
XX 22-OCT-1999; 99US-0160952.
XX 12-OCT-2000; 2000US-0159805.
XX
XX (CURA-) CURAGEN CORP.
XX
XX Prayaga SK, Taupier RJ, Bandaru R;
XX WPI; 2001-308489/32.
XX N-PSDB; AAS00689.
XX
XX New isolated polypeptides, NOV 1-3, having identity to
XX thymosin-beta-10, ephrin type-A receptor 8 and proteoglycans, and
XX polynucleotides, useful for treating male infertility, neurological or
XX cardiac disease or rheumatoid arthritis -
XX
XX Claim 1; Page 8; 102pp; English.

us-09-915-178-2.rag

Wed Jun 4 11:04:59 2003

WPI: 2002-452341/48.
N-PSDB; AAD38157.

DR DR Novel membrane bound and secreted NOV polypeptides, for treating,
XX XX diagnosing and preventing male infertility, neurological, cardiac and
PT PT vascular pathologies, and inflammatory diseases e.g. rheumatoid
PT PT arthritis -

PS Claim 1; Page 12; 180pp; English.

XX The present invention relates to novel membrane bound and secreted NOV
XX proteins and polynucleotides encoding such proteins. Sequences of the
CC invention are useful for treating or preventing NOV-associated disorders
CC in humans and for manufacturing a medicament for treating a syndrome
CC associated with human disease. They are useful for determining the
CC presence of or predisposition to lung cancer. NOV1 compounds are useful
CC for development, differentiation and activation of thymic immune cells,
CC pathologies related to spermatogenesis and male infertility. NOV1 nucleic
CC acids are useful for detecting specific cell types and as specific
CC marker for cancers in tissues. NOV2 and NOV4 compounds are useful to
CC direct the development of nervous system and angiogenesis and for
CC treating neurological, cardiac and vascular pathologies. NOV3 and NOV5
CC compounds are useful for treating various orthopaedic disorders and/or
CC injuries, inflammatory diseases of connective tissue e.g. rheumatoid
CC arthritis, congenital muscular dystrophies, various muscular disorders,
CC fixed deformities (arthrogryposis) and abnormal white matter. They are
CC useful for treating atopy, dysprothrombinaemia, hyperparathyroidism, Leigh
CC centrocyclic lymphoma, cervical carcinoma, hyperparathyroidism, leukæmia,
CC syndrome, hypokalaemic period paralysis, acute promyelocytic leukaemia,
CC NUNA/KARA type, macular dystrophy, vitelliform type, Mcardle myeloma,
CC type 2 Meckel syndrome, multiple endocrine neoplasia I, multiple myeloma,
CC parathyroid adenomatosis 1, prolactinoma, hyperparathyroidism, carcinoid
CC syndrome, digenic retinitis pigmentosa, somatostatinoma, neovascular
CC inflammatory vitreoretinopathy, arthritis and tendonitis. Sequences of
CC the invention are also used in gene therapy and as vaccines. The present
CC sequence is thymosin-beta-10-like (NOV1) protein.

XX Sequence 58 AA;

XX Query Match 60.1%; Score 128; DB 23; Length 58;

XX Best Local Similarity 71.1%; Pred. No. 1.1e-08;

XX Matches 27; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

OY 1 MAHKLDLEETASLDKAKLKATKEMQKNTLMKTTTQEK 38

DB 1 MADKPDIGELASFNKAKLKTKEMQKNTLMKTTTQEK 38

RESULT 8

AAR27109

ID AAR27109 standard; peptide: 43 AA.

XX AAR27109;

XX 26-FEB-1993 (first entry)

XX N-terminus of thymosin beta peptide.

XX Neuropathic; encephalopathy; cerebro-degenerative; neuro-;

XX degenerative.

XX Synthetic.

XX JP04234325-A.

XX 24-AUG-1992.

XX 27-DEC-1990; 90JP-0418312.

XX 27-DEC-1990; 90JP-0418312.

XX PR

XX The sequence represents a thymosin-beta-10-like protein (NOV1).

CC Thymosin-beta-10-like protein, ephrin type-A receptor 8-like protein
CC (NOV2) and proteoglycan-like proteins (NOV3) may be used in the
CC diagnosis, treatment and prevention of disorders caused by abnormal
CC expression or activity of thymosin-beta-10, ephrin type-A receptor 8 and
CC proteoglycans such as fibronectin and fibronectin. The polypeptides of
CC the invention are useful in screening for agents that modulate their
CC activity, and in determining predispositions to disorders. NOV1 is useful
CC for treating conditions involving development, differentiation, and
CC activation of thymic immune cells, in pathologies of neoplasias, in
CC spermatogenesis and male infertility, cells or platelets, in detection of
CC diseases or pathologies of red blood cells can be combined in
CC small cell lung cancer. NOV1 nucleic acids can be combined in
CC chemo-immunotherapeutic anti-cancer treatments. NOV2 is useful for
CC detecting cells expressing GPI-anchored ephrin-A ligands, as a marker for
CC prostate cancer, and in treating neurological, cardiac and vascular
CC disorders. NOV3 (proteoglycan) nucleic acids and proteins are useful for
CC treating orthopaedic disorders and/or injuries, and inflammatory diseases
CC of connective tissues e.g. rheumatoid arthritis, congenital muscular
CC dystrophies.

XX Sequence 58 AA;

XX Query Match 60.1%; Score 128; DB 22; Length 58;

XX Best Local Similarity 71.1%; Pred. No. 1.1e-08;

XX Matches 27; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

OY 1 MAHKLDLEETASLDKAKLKATKEMQKNTLMKTTTQEK 38

DB 1 MADKPDIGELASFNKAKLKTKEMQKNTLMKTTTQEK 38

RESULT 7

AAR23797

ID AAR23797 standard; Protein: 58 AA.

XX AAR23797;

XX 10-SEP-2002 (first entry)

XX Thymosin-beta-10-like (NOV1) protein.

XX Membrane bound protein; secreted NOV protein; spermatogenesis; neoplasia;
XX male infertility; angiogenesis; vascular pathology; orthopaedic disorder;
XX inflammatory disease; congenital muscular dystrophy; muscular disorder;
XX rheumatoid arthritis; fixed deformity; dysprothrombinaemia; cancer;
XX arthrogryposis; hypoprothrombinaemia; hypokalaemic period paralysis;
XX Smith-Lemli-Opitz syndrome; carcinoid tumour; centrocyclic lymphoma;
XX hyperparathyroidism; Leigh syndrome; cervical carcinoma; leukæmia;
XX macular dystrophy; vitelliform type; Mcardle disease; Meckel syndrome;
XX multiple endocrine neoplasia I; multiple myeloma; hyperparathyroidism;
XX parathyroid adenomatosis 1; prolactinoma; digenic retinitis pigmentosa;
XX somatotrophinoma; neovascular inflammatory vitreoretinopathy; arthritis;
XX carcinoid syndrome; atopy; tendonitis; gene therapy; vaccine;
XX thymosin-beta-10-like protein.

XX Unidentified.

XX WO200230979-A2.

XX 18-APR-2002.

XX 10-OCT-2001; 2001WO-US31498.

XX 12-OCT-2000; 2000US-0689486.

XX 13-OCT-2000; 2000US-0687276.

XX 09-OCT-2001; 2001US-0973424.

XX (CURA-) CURAGEN CORP.

XX Prayaga SK, Taupier RJ, Bandaru R;

XX PI

XX (FARH) HOECHST JAPAN LTD.
 XX WPI; 1992-327505/40.
 XX
 XX Neuropathic agent containing thymosin B4, B8, B10 or B11
 XX peptide(s) - for treatment and prevention of cerebro- and
 XX neuro-degenerative diseases
 XX
 XX Claim 1; Page 6; 7pp; Japanese.
 XX
 XX The thymosin beta peptide was used in an agent for the treatment
 XX of encephalopathy and neuropathy. The agent may be used for the
 XX treatment or prevention of cerebro- or neuro-degenerative diseases.
 XX See also AAR27106-10.
 XX
 XX Sequence 43 AA;
 XX
 XX Query Match
 XX Best Local Similarity 59.9%; Score 127.5; DB 13; Length 43;
 XX Matches 31; Conservative 2; Mismatches 9; Indels 1; Gaps 1;
 QY 2 AHKLDLEETASLDKAKLKATEMO-KNTLMTKETTEQKWEIS 43
 Db 1 ADKPDMEETASLDKAKLKATEMO-KNTLMTKETTEQKWEIS 43
 RESULT 9
 AAR96935
 ID AAR96935 standard; protein; 41 AA.
 AC AAR96935;
 XX
 XX 17-OCT-1996 (first entry)
 DT Thymosin beta 13.
 DE
 XX Thymosin beta 4; OAD; obstructive airway disease; sputum;
 KW viscoelasticity; respiratory disorder; respiratory distress syndrome;
 KW chronic bronchitis; asthma; emphysema; cystic fibrosis; actin;
 XX polymerisation.
 OS Synthetic.
 XX
 XX WO9611016-A1.
 PN
 XX 18-APR-1996.
 PD
 XX 06-OCT-1995; 95WO-US12543.
 PF
 XX 07-JUN-1995; 95US-0483572.
 PR 06-OCT-1994; 94US-0318994.
 XX
 XX (ALPH-) ALPHA 1 BIOMEDICALS INC.
 XX (UYWA-) UNIV WASHINGTON MEDICAL CENT GEORGE.
 PI Bauman C, Berman ML, Crockford DR, Goldstein AL;
 PI Kater A, Rubin BK;
 XX WPI; 1996-209656/21.
 XX
 XX Reduction of viscoelasticity of sputum of obstructive airway disease
 XX - for use in treating, e.g. cystic fibrosis.
 XX Disclosure; Figure 1C; 45pp; English.
 XX
 XX Thymosin beta 4 may be used in a composition alongside DNase1 to
 XX reduce the viscoelasticity of sputum causing obstructive airway
 XX disease (OAD). Such disease includes obstructive airway
 XX chronic respiratory distress syndrome, chronic bronchitis, asthma,
 XX emphysema and cystic fibrosis associated with excess actin,
 XX polymerisation (polymerisation of G-actin into F-actin).
 XX

SQ Sequence 41 AA;
 XX
 XX Query Match
 XX Best Local Similarity 54.7%; Score 116.5; DB 17; Length 41;
 XX Matches 28; Conservative 1; Mismatches 8; Indels 1; Gaps 1;
 QY 2 AHKLDLEETASLDKAKLKATEMO-KNTLMTKETTEQK 38
 Db 1 ADKPDMEETASLDKAKLKATEMO-KNTLMTKETTEQK 38
 RESULT 10
 AAY80270
 ID AAY80270 standard; peptide; 41 AA.
 XX
 XX AAY80270;
 AC
 XX 26-MAY-2000 (first entry)
 DT Thymosin beta 4 peptide isoform Thetal3.
 DE
 XX Thymosin beta 4; Tbeta4; tissue repair; angiogenesis; wound repair;
 KW cell migration; vulnary; antiatherosclerotic; cytostatic; diagnosis;
 KW healing disorder; fibrotic disease; ischaemia; atherosclerosis;
 KW cell proliferative disorder.
 XX Unidentified.
 OS
 XX WO200006190-A1.
 PN
 XX 10-FEB-2000.
 PD
 XX 29-JUL-1999; 99WO-US17282.
 PF
 XX 30-JUL-1998; 98US-0094690.
 PR (USSH) US DEPT HEALTH & HUMAN SERVICES.
 XX Kleinman HK, Goldstein A, Malinda KM, Sosne G;
 XX WPI; 2000-195195/17.
 XX
 XX A novel method to promote wound healing comprises administering
 XX thymosin beta 4, useful for promoting tissue repair, angiogenesis and
 XX cell migration
 XX Claim 10; Fig 11; 61pp; English.
 XX
 XX A method has been developed for promoting wound healing in a subject
 XX comprising administering a composition containing a wound healing
 XX polypeptide comprising the amino acid sequence Leu-Lys-Lys-Thr-Glu-Thr
 XX and conservative variants, which have wound healing activity. The wound
 XX healing polypeptide is preferably thymosin beta 4 (Tbeta4) or an isoform
 XX of Tbeta4. The polypeptide has vulnary, antiatherosclerotic and
 XX cytotatic activities. It is useful for promoting wound healing, in vivo
 XX in a subject, especially humans, or ex vivo in tissues (e.g. uro-genital,
 XX gastro-intestinal, cardiovascular, muscle, connective and neural,
 XX especially skin or eye tissue). Antibodies against Tbeta4 can be used
 XX to inhibit/ameliorate wound healing. The level of Tbeta4 in a subject
 XX can be used to diagnose a pathological state characterised by a wound
 XX healing disorder, such as fibrotic disease, ischaemia, atherosclerosis
 XX and cell proliferative disorders. Tbeta4 or its isoforms can also be
 XX used to promote epithelial cell (e.g. skin cells, especially
 XX keratinocytes, corneal cells) migration and angiogenesis. The polypeptide
 XX is able to promote wound repair without any side effects, unlike prior
 XX art. The polypeptide represents a thymosin beta 4 peptide isoform from the present
 XX invention.
 XX
 XX Sequence 41 AA;
 XX
 XX Query Match
 XX Best Local Similarity 54.7%; Score 116.5; DB 21; Length 41;
 XX Matches 28; Conservative 1; Mismatches 8; Indels 1; Gaps 1;

XX

Thymosin beta 4 may be used in a composition alongside DNaseI to reduce the viscoelasticity of sputum causing obstructive airway

A method has been developed for promoting wound healing in a subject comprising administering a composition containing a wound healing polypeptide comprising the amino acid sequence Leu-Lys-Lys-Thr-Glu-Thr and conservative variants, which have wound healing activity. The wound healing polypeptide is preferably thymosin beta 4 (Tbet4) or an isoform of Tbet4. The polypeptide has vulnerary, antiatherosclerotic and cytostatic activities. It is useful for promoting wound healing, *in vivo* in a subject, especially humans, or *ex vivo* in tissues (e.g. urogenital, gastro-intestinal, cardiovascular, muscle, connective and neural, especially skin or eye tissue). Antibodies against Tbet4 can be used to inhibit/ameliorate wound healing. The level of Tbet4 in a subject can be used to diagnose a pathological state characterised by a wound healing disorder, such as fibrotic disease, ischaemia, atherosclerosis and cell proliferative disorders. Tbet4 or its isoforms can also be used to promote epithelial cell (e.g. skin cells, especially keratinocytes, corneal cells) migration and angiogenesis. The polypeptide is able to promote wound repair without any side effects, unlike prior art growth and angiogenic factors used for such treatment. The present sequence represents a thymosin beta 4 peptide isoform from the present invention.

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 2, 2003, 15:01:53 ; Search time 1074 Seconds
(without alignments)
3576.881 Million cell updates/sec

Title: US-09-915-178-1
Perfect score: 132
Sequence: 1 atggcacacaactagaact.....adtgagtgaaatttctga 132

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 2054640 seqs, 14551402878 residues

Word size : 10 *or more*

Total number of hits satisfying chosen parameters: 595492

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database :

GenEmbl.*

- 1: gb_ba.*
- 2: gb_hgt.*
- 3: gb_in.*
- 4: gb_in.*
- 5: gb_ov.*
- 6: gb_pat.*
- 7: gb_ph.*
- 8: gb_pi.*
- 9: gb_pi.*
- 10: gb_ro.*
- 11: gb_sts.*
- 12: gb_sy.*
- 13: gb_un.*
- 14: gb_vl.*
- 15: em_ba.*
- 16: em_fun.*
- 17: em_hum.*
- 18: em_in.*
- 19: em_mu.*
- 20: em_mu.*
- 21: em_or.*
- 22: em_ov.*
- 23: em_pat.*
- 24: em_ph.*
- 25: em_pi.*
- 26: em_ro.*
- 27: em_sts.*
- 28: em_un.*
- 29: em_vl.*
- 30: em_hgt_hum.*
- 31: em_hgt_inv.*
- 32: em_hgt_other.*
- 33: em_hgt_mus.*
- 34: em_hgt_pln.*
- 35: em_hgt_rod.*
- 36: em_hgt_mam.*
- 37: em_hgt_vrt.*
- 38: em_sy.*
- 39: em_hgtgo_hum.*
- 40: em_hgtgo_mus.*
- 41: em_hgtgo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query			DB	ID	Description
		Match	Length	%			
1	132	100.0	147	6	AX399230	Sequence	AX399230 Sequence
2	132	100.0	147	6	AX399232	Sequence	AX399232 Sequence
3	132	100.0	147	6	AX399234	Sequence	AX399234 Sequence
4	132	100.0	170536	9	HS18K17	Homo sapi	AL513423 Homo sapi
5	25	18.9	86594	2	AL645728	Homo sapi	AL645728 Homo sapi
6	25	18.9	133135	2	AL713970	Homo sapi	AL713970 Homo sapi
7	25	18.9	151840	2	AL691446	Mus muscu	AL691446 Mus muscu
8	25	18.9	177632	2	AC093615	Homo sapi	AC093615 Homo sapi
9	22	16.7	60372	2	AC097209	Rattus no	AC097209 Rattus no
10	22	16.7	192314	2	AC098254	Rattus no	AC098254 Rattus no
11	21	15.9	198015	2	AC130181	Rattus no	AC130181 Rattus no
12	20	15.2	38399	9	AP001477	Homo sapi	AP001477 Homo sapi
13	20	15.2	53707	9	AC116359	Homo sapi	AC116359 Homo sapi
14	20	15.2	170595	2	AC025937	Homo sapi	AC025937 Homo sapi
15	20	15.2	187097	2	AC113727	Rattus no	AC113727 Rattus no
16	20	15.2	192202	2	AC094105	Homo sapi	AC094105 Homo sapi
17	20	15.2	204080	2	AC093570	Mus muscu	AC093570 Mus muscu
18	20	15.2	207528	2	AL844896	Mus muscu	AL844896 Mus muscu
19	20	15.2	241585	2	AL844536	Mus muscu	AL844536 Mus muscu
20	20	15.2	263050	1	AP000981	Sulfolobu	AP000981 Sulfolobu
21	20	15.2	340000	9	AP001760	Homo sapi	AP001760 Homo sapi
22	19	14.4	11258	1	AE013648	Yersinia	AE013648 Yersinia
23	19	14.4	12292	1	AE011990	Xanthomon	AE011990 Xanthomon
24	19	14.4	67972	2	AC100719	Mus muscu	AC100719 Mus muscu
25	19	14.4	113654	9	AL606526	Human DNA	AL606526 Human DNA
26	19	14.4	121063	2	AC079482	Mus muscu	AC079482 Mus muscu
27	19	14.4	121063	2	AC079482	Mus muscu	AC079482 Mus muscu
28	19	14.4	127392	9	AC008570	Homo sapi	AC008570 Homo sapi
29	19	14.4	137782	10	AL645686	Mouse DNA	AL645686 Mouse DNA
30	19	14.4	155726	2	AC125887	Rattus no	AC125887 Rattus no
31	19	14.4	157361	10	AL645923	Mouse DNA	AL645923 Mouse DNA
32	19	14.4	158286	2	AC111338	Rattus no	AC111338 Rattus no
33	19	14.4	159642	2	AC120092	Rattus no	AC120092 Rattus no
34	19	14.4	160334	2	AC079549	Mus muscu	AC079549 Mus muscu
35	19	14.4	163942	9	AC034207	Homo sapi	AC034207 Homo sapi
36	19	14.4	166606	2	AC020685	Homo sapi	AC020685 Homo sapi
37	19	14.4	168766	2	AC126443	Mus muscu	AC126443 Mus muscu
38	19	14.4	172191	2	AL845505	Mus muscu	AL845505 Mus muscu
39	19	14.4	172797	2	AC122629	Rattus no	AC122629 Rattus no
40	19	14.4	174138	9	AC008871	Homo sapi	AC008871 Homo sapi
41	19	14.4	174670	2	AC112933	Mus muscu	AC112933 Mus muscu
42	19	14.4	175037	2	AC118879	Rattus no	AC118879 Rattus no
43	19	14.4	176409	2	AC097955	Rattus no	AC097955 Rattus no
44	19	14.4	176967	2	AC022742	Homo sapi	AC022742 Homo sapi
45	19	14.4	177033	2	AC102783	Mus muscu	AC102783 Mus muscu
46	19	14.4	180019	2	AC018623	Homo sapi	AC018623 Homo sapi
47	19	14.4	184157	9	AC110005	Homo sapi	AC110005 Homo sapi
48	19	14.4	185027	2	AC116851	Mus muscu	AC116851 Mus muscu
49	19	14.4	185828	2	AC116058	Rattus no	AC116058 Rattus no
50	19	14.4	186253	9	AC020892	Homo sapi	AC020892 Homo sapi
51	19	14.4	190543	2	AC097794	Rattus no	AC097794 Rattus no
52	19	14.4	194568	2	AC129614	Rattus no	AC129614 Rattus no
53	19	14.4	203578	2	AC113034	Mus muscu	AC113034 Mus muscu
54	19	14.4	209207	2	AC130718	Mus muscu	AC130718 Mus muscu
55	19	14.4	209299	2	AC111022	Mus muscu	AC111022 Mus muscu
56	19	14.4	223133	2	AL683829	Mus muscu	AL683829 Mus muscu
57	19	14.4	224040	2	AC023174	Mus muscu	AC023174 Mus muscu
58	19	14.4	229978	2	AC130218	Mus muscu	AC130218 Mus muscu
59	19	14.4	235050	1	AJ414158	Yersinia	AJ414158 Yersinia
60	19	14.4	250362	2	AC122256	Mus muscu	AC122256 Mus muscu
61	19	14.4	336512	2	AC126029	Mus muscu	AC126029 Mus muscu
62	18	13.6	174	11	G06175	human STS W	G06175 human STS W
63	18	13.6	220	11	G36127	STS h14a185	G36127 STS h14a185
64	18	13.6	243	6	AX381579	Sequence	AX381579 Sequence
65	18	13.6	353	6	AX381286	Sequence	AX381286 Sequence

66	18	13.6	395	6	AX381269	AX381269 Sequence	139	18	13.6	127361	9	AC007283	AC007283 Homo sapi
67	18	13.6	400	9	H0MTHMBX	M92381 Human thymo	c 140	18	13.6	129510	8	OLJN00047	AL606614 Oryza sat
68	18	13.6	428	6	AX381565	AX381565 Sequence	141	18	13.6	129608	9	AL353685	AL353685 Human DNA
69	18	13.6	429	6	AX381695	AX381695 Sequence	142	18	13.6	130421	2	AC121428	AC121428 Rattus no
70	18	13.6	434	6	AX381689	AX381689 Sequence	c 143	18	13.6	133387	2	AP004061	AP004061 Oryza sat
71	18	13.6	438	6	AX381520	AX381520 Sequence	144	18	13.6	134940	2	AC018939	AC018939 Homo sapi
72	18	13.6	438	6	AX381645	AX381645 Sequence	145	18	13.6	137924	2	AC128092	AC128092 Rattus no
73	18	13.6	439	6	AX381972	AX381972 Sequence	c 146	18	13.6	144048	9	AC013762	AC013762 Homo sapi
74	18	13.6	439	9	H0MTHVB10	M20259 Human thymo	147	18	13.6	144152	3	AC007725	AC007725 Drosophill
75	18	13.6	443	6	AX381721	AX381721 Sequence	c 148	18	13.6	145953	3	AP005474	AP005474 Oryza sat
76	18	13.6	444	10	RATTHVB10	M17698 Rat thymosi	c 149	18	13.6	149015	9	AC107992	AC107992 Homo sapi
77	18	13.6	445	6	AX381789	AX381789 Sequence	c 150	18	13.6	149034	9	AF275948	AF275948 Homo sapi
78	18	13.6	446	6	AX381953	AX381953 Sequence	151	18	13.6	149807	10	AL772386	AL772386 Mouse DNA
79	18	13.6	446	6	AX410512	AX410512 Sequence	c 152	18	13.6	151196	9	AC063919	AC063919 Homo sapi
80	18	13.6	446	10	RATTHYBB	M58405 R.norvegicu	c 153	18	13.6	152622	2	AC131391	AC131391 Homo sapi
81	18	13.6	446	11	G22975	G22975 human STS W	154	18	13.6	153087	2	AC123448	AC123448 Rattus no
82	18	13.6	449	6	AX381184	AX381184 Sequence	c 155	18	13.6	154265	2	AC127915	AC127915 Rattus no
83	18	13.6	451	6	AX381252	AX381252 Sequence	c 156	18	13.6	154318	9	AC026791	AC026791 Homo sapi
84	18	13.6	451	6	AX381615	AX381615 Sequence	c 157	18	13.6	155518	9	AC015528	AC015528 Homo sapi
85	18	13.6	452	4	AF294616	AF294616 Bos tauru	158	18	13.6	155930	2	AC109106	AC109106 Rattus no
86	18	13.6	452	6	AX381410	AX381410 Sequence	c 159	18	13.6	157126	2	AC121182	AC121182 Rattus no
87	18	13.6	453	6	AX381336	AX381336 Sequence	160	18	13.6	159360	2	AC122607	AC122607 Rattus no
88	18	13.6	453	6	AX410373	AX410373 Sequence	c 161	18	13.6	160149	2	AC016380	AC016380 Homo sapi
89	18	13.6	453	9	S54005	S54005 thymosin be	c 162	18	13.6	161095	2	AP005647	AP005647 Oryza sat
90	18	13.6	455	6	AX381335	AX381335 Sequence	c 163	18	13.6	162482	9	AC022740	AC022740 Homo sapi
91	18	13.6	456	6	AX381281	AX381281 Sequence	c 164	18	13.6	162643	2	AC113327	AC113327 Mus muscu
92	18	13.6	465	4	AF506973	AF506973 Equus cab	165	18	13.6	166266	2	AP001146	AP001146 Homo sapi
93	18	13.6	465	6	AX381118	AX381118 Sequence	c 166	18	13.6	166543	2	AC128985	AC128985 Rattus no
94	18	13.6	461	6	AX107107	AX107107 Sequence	c 167	18	13.6	167480	9	HSBA18114	HSBA18114 Human DNA
95	18	13.6	500	9	BC016731	BC016731 Homo sapi	c 168	18	13.6	168448	9	AC009469	AC009469 Homo sapi
96	18	13.6	511	9	BC016025	BC016025 Homo sapi	c 169	18	13.6	169341	2	AC114700	AC114700 Rattus no
97	18	13.6	517	6	AX381936	AX381936 Sequence	c 170	18	13.6	169609	2	AC118128	AC118128 Rattus no
98	18	13.6	539	10	RATTHYBA	M58406 R.norvegicu	c 171	18	13.6	170696	2	AC108564	AC108564 Rattus no
99	18	13.6	559	6	AX381221	Z48496 M.musculus	c 172	18	13.6	171030	9	CNS01D5F	CNS01D5F Human chr
100	18	13.6	559	6	AX381221	AX381221 Sequence	173	18	13.6	171391	2	AC125699	AC125699 Rattus no
101	18	13.6	611	5	XLXTB4P	D10692 Xenopus lae	c 174	18	13.6	173817	9	AC103834	AC103834 Homo sapi
102	18	13.6	693	6	AX381860	AX381860 Sequence	175	18	13.6	175127	2	AC016048	AC016048 Homo sapi
103	18	13.6	1173	5	GU28980	U28980 Gallus gall	176	18	13.6	176983	9	AC019172	AC019172 Homo sapi
104	18	13.6	1262	9	H0MTHMB10	M92383 Homo sapien	177	18	13.6	177826	2	AC016852	AC016852 Homo sapi
105	18	13.6	1402	9	AF090913	AF090913 Homo sapi	c 178	18	13.6	177988	2	AC022760	AC022760 Homo sapi
106	18	13.6	1517	6	AX428883	AX428883 Sequence	c 179	18	13.6	179216	2	AC074016	AC074016 Homo sapi
107	18	13.6	2230	9	AK097283	AK097283 Homo sapi	c 180	18	13.6	179992	2	AC095165	AC095165 Rattus no
108	18	13.6	2460	10	AB040468	AB040468 Rattus no	c 181	18	13.6	180699	2	AC079403	AC079403 Homo sapi
109	18	13.6	2489	3	DMCYCLB	X55542 D.melanogas	182	18	13.6	180993	2	AC102164	AC102164 Mus muscu
110	18	13.6	2612	3	DRCCYCB	M33192 D.melanogas	183	18	13.6	181389	2	AC099496	AC099496 Homo sapi
111	18	13.6	2630	3	AX102682	AX102682 Drosophill	184	18	13.6	182165	9	AC093429	AC093429 Homo sapi
112	18	13.6	4000	3	DME006773	AX006773 Drosophill	185	18	13.6	182183	3	AC007889	AC007889 Drosophill
113	18	13.6	4307	9	HSTYL	X96688 H.sapiens m	c 186	18	13.6	183023	2	AC095955	AC095955 Rattus no
114	18	13.6	4407	5	AX072908	AX072908 Gallus ga	c 187	18	13.6	183438	9	AC007738	AC007738 Homo sapi
115	18	13.6	23775	2	AC018178	AC018178 Drosophill	c 188	18	13.6	183999	6	AX092589	AX092589 Sequence
116	18	13.6	32303	9	AC007059	AC007059 Homo sapi	c 189	18	13.6	184689	2	AC009172	AC009172 Homo sapi
117	18	13.6	33419	10	AL662910	AL662910 Mouse DNA	c 190	18	13.6	185125	2	AC101663	AC101663 Mus muscu
118	18	13.6	39852	9	AC006128	AC006128 Homo sapi	c 191	18	13.6	186120	9	AL390755	AL390755 Human DNA
119	18	13.6	43632	1	SC6A5	AL049485 Streptomy	192	18	13.6	186124	2	AC007598	AC007598 Homo sapi
120	18	13.6	49574	2	AC098338	AC098338 Rattus no	c 193	18	13.6	186886	2	AC016297	AC016297 Homo sapi
121	18	13.6	54513	2	AC103847	AC103847 Homo sapi	194	18	13.6	187006	10	AL591436	AL591436 Mouse DNA
122	18	13.6	54873	2	AC101683	AC101683 Homo sapi	c 195	18	13.6	187364	10	AC012295	AC012295 Mus muscu
123	18	13.6	54958	2	AC103850	AC103850 Homo sapi	c 196	18	13.6	188921	2	AC025122	AC025122 Homo sapi
124	18	13.6	57327	10	AC003018	AC003018 Mus muscu	c 197	18	13.6	191438	3	AC008351	AC008351 Drosophill
125	18	13.6	59469	2	AC110024	AC110024 Homo sapi	c 198	18	13.6	194295	2	AC129454	AC129454 Rattus no
126	18	13.6	60528	2	AC020123	AC020123 Drosophill	c 199	18	13.6	195840	2	CNS0000P	CNS0000P Human chr
127	18	13.6	61709	2	AC113001	AC113001 Mus muscu	c 200	18	13.6	196564	2	AL772179	AL772179 Mus muscu
128	18	13.6	68925	9	AC073196	AC073196 Homo sapi	c 201	18	13.6	196674	2	AC091802	AC091802 Mus muscu
129	18	13.6	90767	2	H0510A06	AL442104 Oryza sat	c 202	18	13.6	198148	2	AC096943	AC096943 Rattus no
130	18	13.6	96128	10	AL663077	AL663077 Mouse DNA	c 203	18	13.6	199308	2	AC097367	AC097367 Mus muscu
131	18	13.6	100751	2	AC114144	AC114144 Rattus no	204	18	13.6	199328	2	AC097362	AC097362 Rattus no
132	18	13.6	102591	9	AC002978	AC002978 Homo sapi	c 205	18	13.6	199827	2	AC026373	AC026373 Homo sapi
133	18	13.6	108408	2	AC022210	AC022210 Homo sapi	c 206	18	13.6	200167	2	AC128881	AC128881 Rattus no
134	18	13.6	108952	2	AC096600	AC096600 Rattus no	c 207	18	13.6	201144	9	AF287262	AF287262 Homo sapi
135	18	13.6	111020	2	AC105721	AC105721 Rattus no	208	18	13.6	202500	2	AC124012	AC124012 Mus muscu
136	18	13.6	112911	9	AC005865	AC005865 Homo sapi	209	18	13.6	210891	2	AL844142	AL844142 Mus muscu
137	18	13.6	121584	2	AC008232	AC008232 Drosophill	c 210	18	13.6	210891	2	AC116025	AC116025 Homo sapi
138	18	13.6	126335	2	AC127982	AC127982 Rattus no	211	18	13.6	212835	2	AC007612	AC007612 Homo sapi

c 212	18	13.6	218158	2	AC096363	AC096363 Rattus no	c 285	17	12.9	67683	9	AC092438	AC092438 Homo sapi
c 213	18	13.6	219017	9	AP000936	AP000936 Homo sapi	c 286	17	12.9	69954	2	AC090603	AC090603 Homo sapi
c 214	18	13.6	220462	2	AC073439	AC073439 Homo sapi	c 287	17	12.9	70394	2	AC023011	AC023011 Homo sapi
c 215	18	13.6	226483	3	AE003693	AE003693 Drosophill	c 288	17	12.9	70820	2	AC017451	AC017451 Drosophill
c 216	18	13.6	228919	2	AC090123	AC090123 Mus muscu	c 289	17	12.9	72756	2	AC018401	AC018401 Homo sapi
c 217	18	13.6	235218	2	AC074208	AC074208 Mus muscu	c 290	17	12.9	73916	9	AP000765	AP000765 Homo sapi
c 218	18	13.6	237619	2	AC073765	AC073765 Mus muscu	c 291	17	12.9	74229	2	AC111107	AC111107 Mus muscu
c 219	18	13.6	251256	2	AC099387	AC099387 Rattus no	c 292	17	12.9	79048	2	AC024976	AC024976 Homo sapi
c 220	18	13.6	263190	2	AC015657	AC015657 Mus muscu	c 293	17	12.9	79313	2	AC106500	AC106500 Rattus no
c 221	18	13.6	302225	3	AE003458	AE003458 Drosophill	c 294	17	12.9	80428	2	AC012631	AC012631 Homo sapi
c 222	18	13.6	347750	1	AP002998	AP002998 Mesorhizo	c 295	17	12.9	80655	8	AP004577	AP004577 Lotus jap
c 223	17	12.9	171	10	S40716	S40716 CD45 (signa	c 296	17	12.9	83655	8	AB025633	AB025633 Arabidops
c 224	17	12.9	430	6	AX119979	AX119979 Sequence	c 297	17	12.9	83857	2	AP004012	AP004012 Oryza sat
c 225	17	12.9	500	4	CFA278004	CFA278004 Canis fam	c 298	17	12.9	83989	8	H0505A02	AL512543 Oryza sat
c 226	17	12.9	570	6	AX250477	AX250477 Homo sapi	c 299	17	12.9	85475	9	AC013585	AC013585 Homo sapi
c 227	17	12.9	691	9	HS3326437	AJ326437 Homo sapi	c 300	17	12.9	85519	9	AC112242	AC112242 Homo sapi
c 228	17	12.9	711	9	HS333683	AJ333683 Homo sapi	c 301	17	12.9	87340	9	AL357373	AL357373 Human DNA
c 229	17	12.9	735	9	AF151893	AF151893 Homo sapi	c 302	17	12.9	88176	2	AL390202	AL390202 Oryza sat
c 230	17	12.9	763	9	BC003540	BC003540 Homo sapi	c 303	17	12.9	8937	2	AC096092	AC096092 Rattus no
c 231	17	12.9	763	9	BC009428	BC009428 Homo sapi	c 304	17	12.9	96411	10	AL663114	AL663114 Mouse DNA
c 232	17	12.9	850	9	HMTGFBRII	D28131 Human mRNA	c 305	17	12.9	96540	8	ATAC012328	AC012328 Arabidops
c 233	17	12.9	911	9	HMTGFBRII	D28131 Human mRNA	c 306	17	12.9	98540	8	ATAC012328	AC012328 Arabidops
c 234	17	12.9	928	9	HMTGFBRII	D28131 Human mRNA	c 307	17	12.9	98540	8	ATAC012328	AC012328 Arabidops
c 235	17	12.9	1077	6	AX121123	AX121123 Sequence	c 308	17	12.9	99325	2	AL139427	AL139427 Homo sapi
c 236	17	12.9	1089	5	AF296377	AF296377 Salmo tru	c 309	17	12.9	99591	2	AC013392	AC013392 Homo sapi
c 237	17	12.9	1089	5	AF296381	AF296381 Salmo tru	c 310	17	12.9	101814	2	AC122087	AC122087 Rattus no
c 238	17	12.9	1099	5	AF296370	AF296370 Oncorhync	c 311	17	12.9	104129	2	AC128003	AC128003 Rattus no
c 239	17	12.9	1185	8	AY086277	AY086277 Arabidops	c 312	17	12.9	104782	9	CNS05FEZ	AL359681 Human chr
c 240	17	12.9	1200	6	AX068683	AX068683 Sequence	c 313	17	12.9	106610	9	AC099679	AC099679 Homo sapi
c 241	17	12.9	1227	8	AY070073	AY070073 Arabidops	c 314	17	12.9	106645	2	AC112845	AC112845 Rattus no
c 242	17	12.9	1284	10	AF251010	AF251010 Rattus no	c 315	17	12.9	106821	2	AC020000	AC020000 Drosophill
c 243	17	12.9	1704	6	AX481435	AX481435 Sequence	c 316	17	12.9	107307	2	AC103269	AC103269 Rattus no
c 244	17	12.9	1918	6	E10743	E10743 cDNA encodi	c 317	17	12.9	107458	9	AC023492	AC023492 Homo sapi
c 245	17	12.9	2090	6	E10743	E10743 cDNA encodi	c 318	17	12.9	107467	9	HS958B3	Z93023 Homo sapien
c 246	17	12.9	2090	6	AR031257	AR031257 Sequence	c 319	17	12.9	108752	10	AF084363	AF084363 Mus muscu
c 247	17	12.9	2090	6	AR094643	AR094643 Sequence	c 320	17	12.9	109183	2	AC104339	AC104339 Homo sapi
c 248	17	12.9	2090	6	AR096458	AR096458 Sequence	c 321	17	12.9	111350	9	ALI57814	ALI57814 Human DNA
c 249	17	12.9	2090	6	AR102454	AR102454 Sequence	c 322	17	12.9	112670	9	ALI589165	ALI589165 Human DNA
c 250	17	12.9	2090	6	E1170218	E1170218 Sequence	c 323	17	12.9	113853	9	ALI37853	ALI37853 Human DNA
c 251	17	12.9	2090	6	E31044	E31044 Method for	c 324	17	12.9	114007	9	HSJ1185K9	HSJ1185K9 Human DNA
c 252	17	12.9	2090	9	HMTGFBRII	M85079 Human TGF-b	c 325	17	12.9	115046	2	AC109955	AC109955 Rattus no
c 253	17	12.9	2090	6	AX451991	AX451991 Sequence	c 326	17	12.9	115279	9	AC112506	AC112506 Homo sapi
c 254	17	12.9	2095	6	I78448	I78448 Sequence 1	c 327	17	12.9	117955	9	AC019052	AC019052 Homo sapi
c 255	17	12.9	2153	5	AF287486	AF287486 Oncorhync	c 328	17	12.9	118030	2	AP000667	AP000667 Homo sapi
c 256	17	12.9	2738	4	CPCRGMPPB	Z23014 C.familiar	c 329	17	12.9	118103	2	AC009817	AC009817 Homo sapi
c 257	17	12.9	2741	4	D0GPD8B	L13262 Canis famil	c 330	17	12.9	119308	2	AC094126	AC094126 Rattus no
c 258	17	12.9	2874	3	AY051856	AY051856 Drosophill	c 331	17	12.9	120133	2	AC108334	AC108334 Rattus no
c 259	17	12.9	3000	8	TIN251864	AJ251864 Polyocia	c 332	17	12.9	120846	9	AC018356	AC018356 Homo sapi
c 260	17	12.9	3653	5	AF387816	AF387816 Xenopus l	c 333	17	12.9	123246	8	AP004266	AP004266 Oryza sat
c 261	17	12.9	4592	10	RNU67309	U67309 Rattus norv	c 334	17	12.9	124699	9	AL590617	AL590617 Human DNA
c 262	17	12.9	5057	6	AR054007	AR054007 Sequence	c 335	17	12.9	125291	9	AP000646	AP000646 Homo sapi
c 263	17	12.9	5057	6	AR146183	AR146183 Sequence	c 336	17	12.9	126000	2	AC109739	AC109739 Rattus no
c 264	17	12.9	5057	10	RRBNO5	XS9949 R.rattus MR	c 337	17	12.9	129175	9	AC022103	AC022103 Homo sapi
c 265	17	12.9	5666	9	AB061824	AB061824 Homo sapi	c 338	17	12.9	129380	2	AC126831	AC126831 Rattus no
c 266	17	12.9	5759	6	AX334519	AX334519 Sequence	c 339	17	12.9	130267	2	AC010597	AC010597 Homo sapi
c 267	17	12.9	5759	6	D50683	D50683 Homo sapien	c 340	17	12.9	130912	2	AC119050	AC119050 Gallus ga
c 268	17	12.9	10272	1	AE005682	AE005682 Caulobact	c 341	17	12.9	132870	2	AC119629	AC119629 Rattus no
c 269	17	12.9	12315	3	AF304204	AF304204 Strongylo	c 342	17	12.9	133747	2	AC123129	AC123129 Rattus no
c 270	17	12.9	37182	9	HSU9D4	Z68330 Human DNA s	c 343	17	12.9	133863	2	AC111322	AC111322 Rattus no
c 271	17	12.9	38786	5	AF170972	AF170972 Agelaius	c 344	17	12.9	134506	9	AC005738	AC005738 Homo sapi
c 272	17	12.9	43041	9	U73636	U73636 Human Chrom	c 345	17	12.9	135915	2	AC091028	AC091028 Homo sapi
c 273	17	12.9	45375	5	AF328738	AF328738 Agelaius	c 346	17	12.9	139487	2	AC123355	AC123355 Rattus no
c 274	17	12.9	48385	9	AC118990	AC118990 Homo sapi	c 347	17	12.9	139814	2	AP003964	AP003964 Oryza sat
c 275	17	12.9	50603	2	AC027167	AC027167 Homo sapi	c 348	17	12.9	140000	9	AP001791	AP001791 Homo sapi
c 276	17	12.9	50768	3	AC024853	AC024853 Caenorhab	c 349	17	12.9	141799	2	AC023581	AC023581 Homo sapi
c 277	17	12.9	55777	2	AC115729	AC115729 Mus muscu	c 350	17	12.9	142481	2	AC124332	AC124332 Rattus no
c 278	17	12.9	55946	9	AC104654	AC104654 Homo sapi	c 351	17	12.9	142718	2	AC109663	AC109663 Rattus no
c 279	17	12.9	58133	9	AL357057	AL357057 Human DNA	c 352	17	12.9	143039	2	AC125869	AC125869 Rattus no
c 280	17	12.9	59580	9	AC106875	AC106875 Homo sapi	c 353	17	12.9	145089	9	AC022418	AC022418 Homo sapi
c 281	17	12.9	62156	2	AC117678	AC117678 Mus muscu	c 354	17	12.9	145253	9	AC006329	AC006329 Homo sapi
c 282	17	12.9	63173	2	AC103995	AC103995 Homo sapi	c 355	17	12.9	145737	2	AC105649	AC105649 Rattus no
c 283	17	12.9	63436	2	AC023336	AC023336 Homo sapi	c 356	17	12.9	145910	8	OSG00027	AL606600 Oryza sat
c 284	17	12.9	63479	2	AC124270	AC124270 Homo sapi	c 357	17	12.9	146040	2	AC129734	AC129734 Rattus no
c 285	17	12.9	65782	2	AC096277	AC096277 Rattus no							
c 286	17	12.9	65782	2	AC096277	AC096277 Rattus no							
c 287	17	12.9	65782	2	AC096277	AC096277 Rattus no							
c 288	17	12.9	65782	2	AC096277	AC096277 Rattus no							
c 289	17	12.9	65782	2	AC096277	AC096277 Rattus no							
c 290	17	12.9	65782	2	AC096277	AC096277 Rattus no							
c 291	17	12.9	65782	2	AC096277	AC096277 Rattus no							
c 292	17	12.9	65782	2	AC096277	AC096277 Rattus no							
c 293	17	12.9	65782	2	AC096277	AC096277 Rattus no							
c 294	17	12.9	65782	2	AC096277	AC096277 Rattus no							
c 295	17	12.9	65782	2	AC096277	AC096277 Rattus no							
c 296	17	12.9	65782	2	AC096277	AC096277 Rattus no							
c 297	17	12.9	65782	2	AC096277	AC096277 Rattus no							
c 298	17	12.9	65782	2	AC096277	AC096277 Rattus no							
c 299	17	12.9	65782	2	AC096277	AC096277 Rattus no							
c 300	17	12.9	65782	2	AC096277	AC096277 Rattus no							
c 301	17	12.9	65782	2	AC096277	AC096277 Rattus no							
c 302	17	12.9	65782	2	AC096277	AC096277 Rattus no							
c 303	17	12.9	65782	2	AC096277	AC096277 Rattus no							
c 304	17	12.9	65782	2	AC096277	AC096277 Rattus no							
c 305	17	12.9	65782	2	AC096277	AC096277 Rattus no							

358	17	12.9 147274	2	AP005588	AP005588 Oryza sat	431	17	12.9 174031	4	AC091507	AC091507 Sus scrof
359	17	12.9 148151	9	AL513533	AL513533 Human DNA	C 432	17	12.9 174124	2	AC100778	AC100778 Homo sapi
360	17	12.9 149913	2	AC1010330	AC1010330 Homo sapi	C 433	17	12.9 175053	2	AC023226	AC023226 Homo sapi
361	17	12.9 149963	2	AC125673	AC125673 Rattus no	C 434	17	12.9 175456	2	AC115183	AC115183 Rattus no
C 362	17	12.9 150670	2	AC094577	AC094577 Rattus no	C 435	17	12.9 176209	9	AC009518	AC009518 Homo sapi
363	17	12.9 150885	2	AC023363	AC023363 Homo sapi	C 436	17	12.9 176350	2	AC094034	AC094034 Rattus no
C 364	17	12.9 151340	2	AC020961	AC020961 Mus muscu	C 437	17	12.9 176418	9	AC090811	AC090811 Homo sapi
365	17	12.9 151544	2	AC021650	AC021650 Homo sapi	C 438	17	12.9 176554	2	AC112308	AC112308 Rattus no
366	17	12.9 153500	2	AC128380	AC128380 Rattus no	C 439	17	12.9 176846	2	AC097730	AC097730 Mus muscu
C 367	17	12.9 153763	2	AC114420	AC114420 Mus muscu	C 440	17	12.9 176879	2	AC097730	AC097730 Rattus no
C 368	17	12.9 154109	2	AC108965	AC108965 Rattus no	C 441	17	12.9 177108	2	AC110561	AC110561 Mus muscu
C 369	17	12.9 154236	2	AC123394	AC123394 Rattus no	C 442	17	12.9 177158	2	AC121433	AC121433 Rattus no
C 370	17	12.9 154361	2	AC084189	AC084189 Homo sapi	C 443	17	12.9 177351	2	AC091170	AC091170 Homo sapi
C 371	17	12.9 154723	2	CNS01002	AL1133124 Homo sapi	C 444	17	12.9 178438	9	AC096774	AC096774 Homo sapi
C 372	17	12.9 154728	2	AC098631	AC098631 Rattus no	C 445	17	12.9 178508	2	AC023040	AC023040 Homo sapi
C 373	17	12.9 154758	2	AC016232	AC016232 Homo sapi	C 446	17	12.9 180008	2	AC022962	AC022962 Homo sapi
C 374	17	12.9 155450	2	AC079984	AC079984 Homo sapi	C 447	17	12.9 180421	2	AL805970	AL805970 Mus muscu
C 375	17	12.9 155877	2	AC073631	AC073631 Rattus no	C 448	17	12.9 180423	9	AC012152	AC012152 Homo sapi
C 376	17	12.9 156325	2	AC117926	AC117926 Homo sapi	C 449	17	12.9 180446	2	AP001997	AP001997 Homo sapi
C 377	17	12.9 156491	9	AL591720	AL591720 Human DNA	C 450	17	12.9 180568	2	AC125060	AC125060 Mus muscu
C 378	17	12.9 157536	2	AC015639	AC015639 Rattus no	C 451	17	12.9 180835	2	AC120484	AC120484 Rattus no
C 379	17	12.9 157603	2	AL360011	AL360011 Human DNA	C 452	17	12.9 181141	3	AC008346	AC008346 Drosophil
C 380	17	12.9 157636	9	AL592112	AL592112 Mouse DNA	C 453	17	12.9 181366	2	AC099442	AC099442 Rattus no
C 381	17	12.9 158307	10	AC096631	AC096631 Homo sapi	C 454	17	12.9 181418	2	AC102578	AC102578 Mus muscu
C 382	17	12.9 159322	9	AP001318	AP001318 Homo sapi	C 455	17	12.9 181501	2	AC016005	AC016005 Homo sapi
C 383	17	12.9 159569	9	AC010451	AC010451 Homo sapi	C 456	17	12.9 181581	2	AC113917	AC113917 Rattus no
C 384	17	12.9 159862	9	AL138815	AL138815 Human DNA	C 457	17	12.9 181614	2	AC119820	AC119820 Mus muscu
C 385	17	12.9 160210	9	AL138815	AL138815 Human DNA	C 458	17	12.9 181651	2	AC129351	AC129351 Homo sapi
C 386	17	12.9 160655	2	AC064874	AC064874 Homo sapi	C 459	17	12.9 181732	2	AC127292	AC127292 Mus muscu
C 387	17	12.9 160737	2	AL157816	AL157816 Homo sapi	C 460	17	12.9 182326	2	AC125693	AC125693 Rattus no
C 388	17	12.9 160853	9	AP003788	AP003788 Homo sapi	C 461	17	12.9 182565	2	AC101990	AC101990 Mus muscu
C 389	17	12.9 161289	9	AP000676	AP000676 Homo sapi	C 462	17	12.9 182898	2	AC112691	AC112691 Mus muscu
C 390	17	12.9 161548	9	AP001929	AP001929 Homo sapi	C 463	17	12.9 183244	3	AC027188	AC027188 Homo sapi
C 391	17	12.9 162018	9	AC092610	AC092610 Homo sapi	C 464	17	12.9 183333	10	AL671987	AL671987 Mouse DNA
C 392	17	12.9 162172	2	AC027384	AC027384 Homo sapi	C 465	17	12.9 184466	2	AC113275	AC113275 Mus muscu
C 393	17	12.9 162301	2	AC011023	AC011023 Homo sapi	C 466	17	12.9 184799	2	AC037448	AC037448 Homo sapi
C 394	17	12.9 162397	2	AC0112536	AC0112536 Homo sapi	C 467	17	12.9 185022	10	AL627123	AL627123 Mouse DNA
C 395	17	12.9 162485	2	AC011321	AC011321 Homo sapi	C 468	17	12.9 185108	2	AC118835	AC118835 Rattus no
C 396	17	12.9 162609	9	AC108022	AC108022 Homo sapi	C 469	17	12.9 185324	2	AC026265	AC026265 Homo sapi
C 397	17	12.9 162891	10	AC003061	AC003061 Mouse Chr	C 470	17	12.9 185505	2	AC128731	AC128731 Rattus no
C 398	17	12.9 162971	2	AC094180	AC094180 Rattus no	C 471	17	12.9 185668	3	AC008097	AC008097 Drosophil
C 399	17	12.9 163732	2	AC115395	AC115395 Rattus no	C 472	17	12.9 185713	9	CNS010107	AL138499 Human chr
C 400	17	12.9 163995	2	AC095112	AC095112 Rattus no	C 473	17	12.9 185822	2	AC073554	AC073554 Homo sapi
C 401	17	12.9 164598	2	AC093212	AC093212 Homo sapi	C 474	17	12.9 186105	2	AC010784	AC010784 Homo sapi
C 402	17	12.9 165240	2	AC011040	AC011040 Homo sapi	C 475	17	12.9 186373	9	AC006229	AC006229 Homo sapi
C 403	17	12.9 165434	9	AC022448	AC022448 Homo sapi	C 476	17	12.9 186526	10	AL596083	AL596083 Mouse DNA
C 404	17	12.9 165889	10	AL606841	AL606841 Mouse DNA	C 477	17	12.9 186562	2	AL451079	AL451079 Homo sapi
C 405	17	12.9 166634	2	AC092602	AC092602 Homo sapi	C 478	17	12.9 186744	2	AC131561	AC131561 Rattus no
C 406	17	12.9 167692	2	AC128709	AC128709 Homo sapi	C 479	17	12.9 187067	9	AC105389	AC105389 Homo sapi
C 407	17	12.9 167888	2	AC121651	AC121651 Rattus no	C 480	17	12.9 187143	8	OSJN00107	AL606729 Oryza sat
C 408	17	12.9 167937	2	AC095024	AC095024 Sus scrof	C 481	17	12.9 187278	2	AC110895	AC110895 Mus muscu
C 409	17	12.9 168396	2	AL590010	AL590010 Homo sapi	C 482	17	12.9 187428	2	AC009613	AC009613 Homo sapi
C 410	17	12.9 168556	2	AC097677	AC097677 Rattus no	C 483	17	12.9 187492	2	AC129260	AC129260 Rattus no
C 411	17	12.9 168607	2	AC121994	AC121994 Mus muscu	C 484	17	12.9 187710	9	CNS010107	AL135838 Human chr
C 412	17	12.9 168933	2	AC127903	AC127903 Rattus no	C 485	17	12.9 188561	2	AC103201	AC103201 Rattus no
C 413	17	12.9 169758	2	AL713861	AL713861 Mus muscu	C 486	17	12.9 188816	2	AC115271	AC115271 Rattus no
C 414	17	12.9 169781	10	AF336378	AF336378 Mus muscu	C 487	17	12.9 190119	2	AP001814	AP001814 Homo sapi
C 415	17	12.9 170632	9	AL136321	AL136321 Human DNA	C 488	17	12.9 190136	2	AC094982	AC094982 Rattus no
C 416	17	12.9 170639	9	AL391065	AL391065 Human DNA	C 489	17	12.9 190627	10	AL671982	AL671982 Mouse DNA
C 417	17	12.9 170928	9	AL356740	AL356740 Homo sapi	C 490	17	12.9 190842	2	AC109939	AC109939 Rattus no
C 418	17	12.9 171027	2	AP000756	AP000756 Homo sapi	C 491	17	12.9 191403	2	AC105836	AC105836 Mouse DNA
C 419	17	12.9 171279	2	AC036185	AC036185 Homo sapi	C 492	17	12.9 191520	10	AL591598	AL591598 Homo sapi
C 420	17	12.9 171707	2	AC108255	AC108255 Homo sapi	C 493	17	12.9 191540	9	AC010984	AC010984 Homo sapi
C 421	17	12.9 171721	2	AC015744	AC015744 Homo sapi	C 494	17	12.9 191832	30	AC091104	AC091104 Homo sapi
C 422	17	12.9 172203	9	AC005863	AC005863 Homo sapi	C 495	17	12.9 191883	9	AC097469	AC097469 Homo sapi
C 423	17	12.9 172468	2	AP005589	AP005589 Oryza sat	C 496	17	12.9 192042	2	AC063969	AC063969 Mus muscu
C 424	17	12.9 173077	2	AC024511	AC024511 Homo sapi	C 497	17	12.9 192428	2	AC102655	AC102655 Rattus no
C 425	17	12.9 173083	2	AC034272	AC034272 Homo sapi	C 498	17	12.9 192453	2	AC129457	AC129457 Rattus no
C 426	17	12.9 173106	9	AC068985	AC068985 Homo sapi	C 499	17	12.9 192954	2	AC130134	AC130134 Homo sapi
C 427	17	12.9 173201	9	AC008169	AC008169 Homo sapi	C 500	17	12.9 192961	2	AC126183	AC126183 Homo sapi
C 428	17	12.9 173438	2	AC108527	AC108527 Rattus no	C 501	17	12.9 192961	2	AC018886	AC018886 Homo sapi
C 429	17	12.9 173713	2	AC095840	AC095840 Rattus no	C 502	17	12.9 194181	2	AC099694	AC099694 Mus muscu
C 430	17	12.9 173935	10	AL670285	AL670285 Mouse DNA	C 503	17	12.9 194212	2		

504	17	12.9	194377	2	AC021726	AC021726 Homo sapi	577	16	12.1	307	6	AX187986	AX187986 Sequence
505	17	12.9	195386	2	AL805961	AL805961 Homo sapi	578	16	12.1	311	6	AX186432	AX186432 Sequence
506	17	12.9	196097	10	AL714006	AL714006 Mouse DNA	579	16	12.1	331	6	AX185097	AX185097 Sequence
507	17	12.9	196175	10	AC069019	AC069019 Mus muscu	c 580	16	12.1	400	11	GI3730	SHGC-13010
508	17	12.9	196520	2	AC083830	AC083830 Homo sapi	581	16	12.1	406	1	AF455091	AF455091 Unculture
c 509	17	12.9	197088	2	AC107502	AC107502 Rattus no	c 582	16	12.1	429	6	AX329988	AX329988 Sequence
510	17	12.9	197143	2	AC069173	AC069173 Homo sapi	c 583	16	12.1	444	6	AX331994	AX331994 Sequence
511	17	12.9	197946	10	AC098722	AC098722 Mus muscu	c 584	16	12.1	444	6	AX407484	AX407484 Sequence
512	17	12.9	199071	10	AL669938	AL669938 Mouse DNA	c 585	16	12.1	486	11	AU027620	Rattus no
c 513	17	12.9	199101	10	MMAE000665	AE000665 Mus muscu	c 586	16	12.1	528	6	AX385616	AX385616 Sequence
514	17	12.9	200124	2	AC129675	AC129675 Homo sapi	587	16	12.1	576	5	AF212851	AF212851 Aulonocar
c 515	17	12.9	200380	9	AC022898	AC022898 Homo sapi	c 588	16	12.1	799	4	AF325552	AF325552 Bos tauru
c 516	17	12.9	200488	2	AC111048	AC111048 Mus muscu	c 589	16	12.1	834	11	CNS06EQC	AL395434 T7 end of
c 517	17	12.9	200704	9	AC099781	AC099781 Homo sapi	c 590	16	12.1	865	9	HS338652	HS338652 Homo sapi
c 518	17	12.9	200767	2	AC110635	AC110635 Rattus no	c 591	16	12.1	960	10	RATUGT1B4H	D38068 Rattus norv
519	17	12.9	200885	2	AC119856	AC119856 Mus muscu	c 592	16	12.1	969	8	AB001020	AB001020 Schizosac
520	17	12.9	201399	9	AL357149	AL357149 Human DNA	593	16	12.1	1001	3	AY071393	AY071393 Drosophil
521	17	12.9	201412	9	AC096921	AC096921 Homo sapi	594	16	12.1	1018	1	AF526993	AF526993 Unculture
522	17	12.9	201699	9	AC079844	AC079844 Homo sapi	c 595	16	12.1	1028	10	BC002172	BC002172 Mus muscu
523	17	12.9	203111	2	AC096313	AC096313 Rattus no	596	16	12.1	1043	5	AF212849	AF212849 Aulonocar
524	17	12.9	203324	10	AL662780	AL662780 Mouse DNA	c 597	16	12.1	1053	3	TBR315078	TBR315078 Trypanoso
525	17	12.9	203558	2	AC118395	AC118395 Rattus no	598	16	12.1	1094	5	AF212850	AF212850 Aulonocar
c 526	17	12.9	203699	2	AC040973	AC040973 Homo sapi	599	16	12.1	1095	5	XELXPACB2	D31895 Xenopus lae
527	17	12.9	203956	2	AC094828	AC094828 Rattus no	600	16	12.1	1212	6	AX086207	AX086207 Sequence
528	17	12.9	204477	2	AC122374	AC122374 Mus muscu	601	16	12.1	1212	9	HSM801628	HSM801628 Homo sapi
529	17	12.9	204888	2	AL731800	AL731800 Mus muscu	602	16	12.1	1213	3	AY060898	AY060898 Drosophil
c 530	17	12.9	205759	10	AL669853	AL669853 Mouse DNA	603	16	12.1	1238	9	AF136971	AF136971 Homo sapi
c 531	17	12.9	207086	2	AC034290	AC034290 Homo sapi	c 604	16	12.1	1259	8	AY040533	AY040533 Gossypium
c 532	17	12.9	207122	2	AC117235	AC117235 Mus muscu	605	16	12.1	1265	9	BC002444	BC002444 Homo sapi
533	17	12.9	207877	2	AC129206	AC129206 Mus muscu	606	16	12.1	1277	5	XELXPACB	D31894 Xenopus lae
c 534	17	12.9	207922	2	AC073790	AC073790 Mus muscu	c 607	16	12.1	1344	5	XL090895	XL090895 Xenopus lae
c 535	17	12.9	208187	2	AL583884	AL583884 Mus muscu	608	16	12.1	1385	9	AK022861	AK022861 Homo sapi
c 536	17	12.9	208709	9	CNS01RH3	AL592162 Mus muscu	609	16	12.1	1406	14	HSMA41A	HSMA41A Gallid herp
c 537	17	12.9	208953	2	AC095986	AC095986 Rattus no	610	16	12.1	1453	6	AX073680	AX073680 Sequence
c 538	17	12.9	209507	2	AC095986	AC095986 Rattus no	611	16	12.1	1464	9	HSM801555	HSM801555 Homo sapi
c 539	17	12.9	209593	2	AC126279	AC126279 Homo sapi	c 612	16	12.1	1641	8	YSPRGS4	YSPRGS4 Yeast (S.po
c 540	17	12.9	210105	2	AC125948	AC125948 Rattus no	613	16	12.1	1710	9	AF339814	AF339814 Homo sapi
541	17	12.9	211759	2	AC106937	AC106937 Rattus no	614	16	12.1	1815	14	GHU17702	GHU17702 Gallid herp
542	17	12.9	213916	2	AC073473	AC073473 Homo sapi	615	16	12.1	1926	9	BC033843	BC033843 Homo sapi
c 543	17	12.9	214709	2	AC116514	AC116514 Mus muscu	616	16	12.1	1980	8	SCABC1G	SCABC1G S.cerevisia
544	17	12.9	214800	2	AC112851	AC112851 Rattus no	c 617	16	12.1	2199	6	EL7395	EL7395 cDNA encodi
545	17	12.9	215516	2	AC024694	AC024694 Mus muscu	c 618	16	12.1	2199	8	AB010105	AB010105 Grifolia f
c 546	17	12.9	217073	2	AC094162	AC094162 Rattus no	619	16	12.1	2274	9	AF134159	AF134159 Homo sapi
c 547	17	12.9	217186	2	AC118472	AC118472 Mus muscu	620	16	12.1	2335	10	H010058S08	AF229212 Mus muscu
548	17	12.9	218747	2	AL611985	AL611985 Mus muscu	621	16	12.1	2410	3	AF077868	AF077868 Caenorhab
549	17	12.9	220035	2	AC122047	AC122047 Mus muscu	622	16	12.1	2527	9	AK027639	AK027639 Homo sapi
c 550	17	12.9	221883	2	AC104752	AC104752 Mus muscu	623	16	12.1	2545	9	AK023008	AK023008 Homo sapi
551	17	12.9	221898	10	AL669891	AL669891 Mouse DNA	624	16	12.1	2638	9	AK056497	AK056497 Homo sapi
c 552	17	12.9	223496	2	AL732573	AL732573 Mus muscu	625	16	12.1	2639	1	AB081807	AB081807 Streptomy
c 553	17	12.9	225957	9	AC069175	AC069175 Homo sapi	626	16	12.1	2696	9	BC018759	BC018759 Homo sapi
c 554	17	12.9	226467	2	AC121855	AC121855 Mus muscu	627	16	12.1	2741	14	HSMANT1	L29643 Gallid herp
c 555	17	12.9	226576	3	CNS07BGE	AL590447 chromosom	c 628	16	12.1	2817	6	EL7394	EL7394 gDNA encodi
556	17	12.9	227801	2	AL713993	AL713993 Homo sapi	629	16	12.1	2847	8	AY091029	AY091029 Arabidops
557	17	12.9	229957	10	AL670231	AL670231 Mouse DNA	630	16	12.1	2999	9	HSM803265	HSM803265 Homo sapi
c 558	17	12.9	234019	2	AC117951	AC117951 Rattus no	631	16	12.1	3149	6	AX451700	AX451700 Sequence
c 559	17	12.9	238521	2	AC078932	AC078932 Mus muscu	c 632	16	12.1	3240	8	AB010104	AB010104 Grifolia f
c 560	17	12.9	244804	2	AC109236	AC109236 Mus muscu	633	16	12.1	3270	6	AB3858	AB3858 Sequence 3
561	17	12.9	256804	2	AC113622	AC113622 Rattus no	634	16	12.1	3308	14	HEU86698	HEU86698 Human endog
562	17	12.9	260699	2	AC006893	AC006893 Caenorhab	c 635	16	12.1	3392	8	PECXYLAN	PECXYLAN Penicillium
563	17	12.9	267077	2	AL773532	AL773532 Mus muscu	c 636	16	12.1	3424	9	AK055899	AK055899 Homo sapi
564	17	12.9	273744	3	AE003558	AE003558 Drosophil	637	16	12.1	3437	6	AX46720	AX46720 Sequence 3
c 565	17	12.9	281797	3	AE003794	AE003794 Drosophil	638	16	12.1	3437	6	AR110167	AR110167 Sequence
c 566	17	12.9	286552	2	AC079559	AC079559 Mus muscu	639	16	12.1	3437	6	AR211698	AR211698 Sequence
567	17	12.9	332050	1	AP005276	AP005276 Corynebac	c 570	16	12.1	3437	9	HSMPM1	Z46481 H.sapiens m
568	17	12.9	349980	6	AX127146	AX127146 Sequence	640	16	12.1	3437	6	AX333688	AX333688 Sequence
569	16	12.1	215	5	OPU02706	U02706 Oreochromis	641	16	12.1	3437	9	HSMPM1	Z46481 H.sapiens m
c 570	16	12.1	215	5	OPSA005	W5735 O.placidus	642	16	12.1	3456	6	AR110166	AR110166 Sequence
571	16	12.1	220	6	AR069084	AR069084 Sequence	643	16	12.1	3456	6	AR211697	AR211697 Sequence
572	16	12.1	234	5	SGU02720	U02720 Sarcotherodo	644	16	12.1	3513	1	AK095489	AK095489 Homo sapi
573	16	12.1	234	5	SGSATA04	W5727 S.gallilaes	645	16	12.1	3574	1	KPNBUDOPRN	L04507 Klebsellia
574	16	12.1	235	5	SGU02718	U02718 Sarcotherodo	646	16	12.1	3604	9	AK055469	AK055469 Homo sapi
575	16	12.1	235	5	SGSATA02	W5719 S.gallilaes	647	16	12.1	3762	8	AY078948	AY078948 Arabidops
c 576	16	12.1	268	11	G65736	G65736 stcb20F6_47	648	16	12.1	3837	9	AF116574	AF116574 Homo sapi

c 650	16	12.1	3858	1	AF240778	AF240778 Pasteurel	c 723	16	12.1	30213	3	U41032	U41032
c 651	16	12.1	4013	10	MMAR001794	AF0011794 Mus muscu	c 724	16	12.1	32682	2	AC014984	AC014984
c 652	16	12.1	4023	9	SCY350251	AF500251 Homo sapi	c 725	16	12.1	32828	3	CEK08E5	CEK08E5
c 653	16	12.1	4076	8	SCYGL119W	AF50251 S.cerevisia	c 726	16	12.1	33274	8	SPCP20C8	SPCP20C8
c 654	16	12.1	4308	1	PMT0XA	Z72641 Pasteurella	c 727	16	12.1	34195	3	U52003	U52003
c 655	16	12.1	4316	1	PMT0TX	X52478 Pasteurella	c 728	16	12.1	34589	3	AC112240	AC112240
c 656	16	12.1	4316	1	PMPMTX	Z28388 P.multocida	c 729	16	12.1	35251	2	AC113511	AC113511
c 657	16	12.1	4380	1	PMT0X	X51512 Pasteurella	c 730	16	12.1	35782	3	AF039711	AF039711
c 658	16	12.1	4380	6	A08187	X08187 Synthetic n	c 731	16	12.1	36052	3	L14429	L14429
c 659	16	12.1	4380	6	A27814	A27814 P.multicoda	c 732	16	12.1	36133	3	HS007973	HS007973
c 660	16	12.1	4380	6	AR107541	AR107541 Sequence	c 733	16	12.1	36450	2	AC135380	AC135380
c 661	16	12.1	4413	9	HS012755	AJ012755 Homo sapi	c 734	16	12.1	36598	2	AC095907	AC095907
c 662	16	12.1	4428	10	MM081030	U81030 Mus musculu	c 735	16	12.1	36727	9	AL512289	AL512289
c 663	16	12.1	4591	9	AB014534	AB014534 Homo sapi	c 736	16	12.1	38155	6	AR204183	AR204183
c 664	16	12.1	4664	10	AF079520	AF079520 Mus muscu	c 737	16	12.1	38491	9	AC006049	AC006049
c 665	16	12.1	4732	9	AB051522	AB051522 Homo sapi	c 738	16	12.1	38734	8	SPCC338	SPCC338
c 666	16	12.1	4815	6	AX360058	AX360058 Sequence	c 739	16	12.1	39099	9	AF099919	AF099919
c 667	16	12.1	4946	9	HUMPRPH1	M13057 Human acidi	c 740	16	12.1	39359	9	AC000084	AC000084
c 668	16	12.1	4946	11	G28580	G28580 human STS s	c 741	16	12.1	40489	8	SPCC569	SPCC569
c 669	16	12.1	5153	10	AF083372	AF083372 Mus muscu	c 742	16	12.1	41372	9	AL359082	AL359082
c 670	16	12.1	5184	9	AF041835	AF041835 Homo sapi	c 743	16	12.1	41671	9	AC010518	AC010518
c 671	16	12.1	5454	1	AF315586	AF315586 Pseudomon	c 744	16	12.1	41741	9	AL356500	AL356500
c 672	16	12.1	5500	1	PP030471	U30471 Plasmid pSA	c 745	16	12.1	41741	9	AL356500	AL356500
c 673	16	12.1	5792	9	AL713876	AL713876 Human DNA	c 746	16	12.1	42304	9	AC002491	AC002491
c 674	16	12.1	6018	9	S44029	S44029 red visual	c 747	16	12.1	42366	3	CEP25H2	CEP25H2
c 675	16	12.1	6299	9	AB007864	AB007864 Homo sapi	c 748	16	12.1	42886	9	AC016762	AC016762
c 676	16	12.1	6390	6	AX097470	AX097470 Sequence	c 749	16	12.1	43000	9	HS3D19	HS3D19
c 677	16	12.1	6431	1	ECU12441	U12441 Escherichia	c 750	16	12.1	43307	9	HSAC002113	HSAC002113
c 678	16	12.1	6492	8	SJRI8RNA	Z32848 S.japonicus	c 751	16	12.1	43440	3	CEP2005	CEP2005
c 679	16	12.1	6706	10	AF017806	AF017806 Mus muscu	c 752	16	12.1	43473	9	AF022141	AF022141
c 680	16	12.1	6940	6	AX409077	AX409077 Sequence	c 753	16	12.1	43473	9	AF020801	AF020801
c 681	16	12.1	6940	9	D79992	D79992 Human mRNA	c 754	16	12.1	44548	9	AC005862	AC005862
c 682	16	12.1	6972	9	HAAXTRSYV	X90840 H.sapiens m	c 755	16	12.1	45000	5	FRU90882	FRU90882
c 683	16	12.1	7138	9	HSX805540	AL83444 Homo sapi	c 756	16	12.1	46288	5	AL354764	AL354764
c 684	16	12.1	7286	1	AE005332	AE005332 Escherich	c 757	16	12.1	47352	2	AC129582	AC129582
c 685	16	12.1	7865	8	SP18RRNAA	Z19578 S.pombe gen	c 758	16	12.1	47400	2	AC129582	AC129582
c 686	16	12.1	7878	6	AX332512	AX332512 Sequence	c 759	16	12.1	47400	2	AC130262	AC130262
c 687	16	12.1	7878	9	HSTRE210	X63546 H.sapiens m	c 760	16	12.1	48197	2	AC120152	AC120152
c 688	16	12.1	8201	6	I76205	I76205 Sequence 9	c 761	16	12.1	48433	2	AC120152	AC120152
c 689	16	12.1	8201	9	HSTRE213	X63547 H.sapiens m	c 762	16	12.1	48908	6	AR204241	AR204241
c 690	16	12.1	8495	3	AF017096	AF017096 Drosophill	c 763	16	12.1	50170	2	AL355872	AL355872
c 691	16	12.1	9152	9	HSNFX1112	Z47735 H.sapiens N	c 764	16	12.1	50314	2	AC107912	AC107912
c 692	16	12.1	9649	5	AF111198	AF256651 Calman cr	c 765	16	12.1	50314	2	AC100970	AC100970
c 693	16	12.1	9677	9	AF111198	AF111198 Homo sapi	c 766	16	12.1	51173	2	AC103764	AC103764
c 694	16	12.1	10029	1	AE005325	AE005325 Escherich	c 767	16	12.1	52083	2	AC100992	AC100992
c 695	16	12.1	10029	1	AE013918	AE013918 versinia	c 768	16	12.1	53372	2	AC098413	AC098413
c 696	16	12.1	10062	1	AE010425	AE010425 Methanopy	c 769	16	12.1	53721	2	AC087331	AC087331
c 697	16	12.1	10129	1	AE002081	AE002081 Deinococc	c 770	16	12.1	54784	2	AC017732	AC017732
c 698	16	12.1	10260	2	AC129145	AC129145 Rattus no	c 771	16	12.1	55777	2	AC115729	AC115729
c 699	16	12.1	10758	1	AE001231	AE001231 Treponema	c 772	16	12.1	55892	2	AC125102	AC125102
c 700	16	12.1	11129	1	AE001069	AE001069 Archaeogl	c 773	16	12.1	57049	9	HS0385A12	HS0385A12
c 701	16	12.1	11301	3	AF139060	AF139060 Caenorhab	c 774	16	12.1	57578	9	AL589825	AL589825
c 702	16	12.1	11325	1	AE004406	AE004406 Vibrio ch	c 775	16	12.1	57868	2	AC104022	AC104022
c 703	16	12.1	11809	1	AE005290	AE005290 Escherich	c 776	16	12.1	58344	2	AC113788	AC113788
c 704	16	12.1	12848	6	AR204356	AR204356 Sequence	c 777	16	12.1	58511	2	AC131414	AC131414
c 705	16	12.1	13045	14	GH017705	U17705 Gallid herp	c 778	16	12.1	59509	2	AC123720	AC123720
c 706	16	12.1	13104	9	AE004656	AE004656 Pseudomon	c 779	16	12.1	61344	9	AL607076	AL607076
c 707	16	12.1	13236	9	AB012113	AB012113 Homo sapi	c 780	16	12.1	61344	9	AL607076	AL607076
c 708	16	12.1	13505	1	AE005441	AE005441 Escherich	c 781	16	12.1	62072	2	AC101766	AC101766
c 709	16	12.1	13624	1	AE006459	AE006459 Escherich	c 782	16	12.1	62294	2	AC115977	AC115977
c 710	16	12.1	13847	3	HS175B9B	Z69652 Human DNA s	c 783	16	12.1	62974	2	AC130137	AC130137
c 711	16	12.1	14302	3	AF068718	AF068718 Caenorhab	c 784	16	12.1	64196	2	AC116492	AC116492
c 712	16	12.1	17056	6	AR164808	AR164808 Sequence	c 785	16	12.1	64351	2	AC110804	AC110804
c 713	16	12.1	18120	9	HSB20F6	Z82174 Human DNA s	c 786	16	12.1	65028	2	AC111051	AC111051
c 714	16	12.1	18765	1	AE004731	AE004731 Pseudomon	c 787	16	12.1	65028	2	AC127651	AC127651
c 715	16	12.1	20385	1	AE008892	AE008892 Salmonell	c 788	16	12.1	65700	2	AC107370	AC107370
c 716	16	12.1	21480	9	HS0C8B6	Z68193 Human DNA s	c 789	16	12.1	67066	2	AC121141	AC121141
c 717	16	12.1	23106	9	HS0C8B6	AJ271718 Homo sapi	c 790	16	12.1	68166	2	AC084712	AC084712
c 718	16	12.1	23506	9	HS0C8B6	AE008863 Salmonell	c 791	16	12.1	68314	9	AL138798	AL138798
c 719	16	12.1	24453	9	HS0C8B6	AE008863 Salmonell	c 792	16	12.1	68405	2	AC121112	AC121112
c 720	16	12.1	26100	3	HSV434E11	AL022017 Drosophill	c 793	16	12.1	68678	2	AC111027	AC111027
c 721	16	12.1	26881	9	HSV434E11	Z70758 Human DNA s	c 794	16	12.1	68785	5	AL591664	AL591664
c 722	16	12.1	28650	8	SPAC22G7	Z54328 S.pombe chr	c 795	16	12.1	69669	2	AC102217	AC102217
										70873	2	AC119927	AC119927

796	16	12.1	71763	2	AC100350	Mus muscu	869	16	12.1	97600	8	AP004570	Oryza sat
797	16	12.1	71820	2	AC100627	Mus muscu	c 870	16	12.1	97710	9	AL445210	Human DNA
798	16	12.1	7281	2	HSF20148	Homo sapi	c 871	16	12.1	99778	2	AC112883	Rattus no
799	16	12.1	72308	2	AC104994	Homo sapi	c 872	16	12.1	99901	9	AC112715	Homo sapi
800	16	12.1	72511	2	AC025954	Homo sapi	c 873	16	12.1	100259	9	AC005800	Homo sapi
801	16	12.1	73359	2	AC025012	Homo sapi	c 874	16	12.1	100292	10	AL627346	Mouse DNA
802	16	12.1	73746	2	AC101286	Mus muscu	c 875	16	12.1	100300	2	AC114970	Homo sapi
803	16	12.1	73920	10	AC091397	Rattus no	c 876	16	12.1	100746	8	AC005489	Genomic s
804	16	12.1	74229	2	AC111107	Mus muscu	c 877	16	12.1	101055	2	AC107941	Homo sapi
805	16	12.1	74315	2	AC023029	Homo sapi	c 878	16	12.1	101759	2	AP005365	Homo sapi
806	16	12.1	74712	2	AC121278	Mus muscu	c 879	16	12.1	102477	9	AC008935	Homo sapi
807	16	12.1	75367	2	AC107515	Rattus no	c 880	16	12.1	102540	8	AC079281	Arabidops
808	16	12.1	75525	9	AL391561	Human DNA	c 881	16	12.1	102840	9	AL161719	Human DNA
809	16	12.1	76746	2	AC121670	Rattus no	c 882	16	12.1	103122	9	AC005663	Homo sapi
810	16	12.1	76746	2	AC121670	Rattus no	c 883	16	12.1	103243	2	AC106472	Rattus no
811	16	12.1	76966	2	AC032037	Homo sapi	c 884	16	12.1	103458	2	AC127085	Rattus no
812	16	12.1	77290	8	AC074284	Arabidops	c 885	16	12.1	104171	2	AC019499	Drosophil
813	16	12.1	77728	9	AP000146	Homo sapi	c 886	16	12.1	104492	9	AC003985	Homo sapi
814	16	12.1	77728	9	AP000231	Homo sapi	c 887	16	12.1	104733	2	AC122598	Rattus no
815	16	12.1	79054	8	F21F23	Sequence	c 888	16	12.1	104753	2	AC114062	Rattus no
816	16	12.1	79304	2	AC103063	Rattus no	c 889	16	12.1	105250	2	AP004099	Oryza sat
817	16	12.1	79899	8	AB0111474	Arabidops	c 890	16	12.1	105573	2	AC094644	Rattus no
818	16	12.1	80460	10	AL513014	Mouse DNA	c 891	16	12.1	105866	2	AC021602	Homo sapi
819	16	12.1	80828	9	AC093247	Homo sapi	c 892	16	12.1	106606	2	AC095551	Rattus no
820	16	12.1	81044	2	AC103124	Rattus no	c 893	16	12.1	107733	9	AC008472	Homo sapi
821	16	12.1	81054	9	AL356052	Human DNA	c 894	16	12.1	107745	9	AL139042	Human DNA
822	16	12.1	81283	2	AC097851	Rattus no	c 895	16	12.1	108055	2	AC127721	Rattus no
823	16	12.1	81410	9	AP004715	Homo sapi	c 896	16	12.1	108317	2	AC123104	Rattus no
824	16	12.1	81414	8	AB024037	Arabidops	c 897	16	12.1	108523	9	AC000159	Homo sapi
825	16	12.1	81580	8	AP000736	Arabidops	c 898	16	12.1	108742	2	AC023817	Homo sapi
826	16	12.1	82404	9	AL133399	Human DNA	c 899	16	12.1	109219	9	AF159227	Homo sapi
827	16	12.1	82435	9	AC092424	Homo sapi	c 900	16	12.1	109225	2	AC091417	Rattus no
828	16	12.1	82460	2	AC107433	Rattus no	c 901	16	12.1	109299	3	AC117838	Caenorhab
829	16	12.1	83097	8	AB025624	Arabidops	c 902	16	12.1	110000	2	AC009579	0
830	16	12.1	84001	9	AL162739	Human DNA	c 903	16	12.1	110000	2	AC087331	3
831	16	12.1	84657	2	AC129132	Rattus no	c 904	16	12.1	110000	2	AC087331	3
832	16	12.1	84686	9	AL138685	Human DNA	c 905	16	12.1	110000	2	AC117400	2
833	16	12.1	84780	9	AC005885	Homo sapi	c 906	16	12.1	110000	2	AC122688	2
834	16	12.1	84866	2	AC103533	Rattus no	c 907	16	12.1	110000	2	AC123950	1
835	16	12.1	84872	8	AC006593	Arabidops	c 908	16	12.1	110000	2	AC123950	2
836	16	12.1	85472	2	AC095888	Rattus no	c 909	16	12.1	110000	2	AC125140	3
837	16	12.1	85579	9	AC004744	Homo sapi	c 910	16	12.1	110000	2	HSY313F4	2
838	16	12.1	85750	2	AC103346	Rattus no	c 911	16	12.1	110000	2	AC024562	0
839	16	12.1	86371	9	AL355672	Human DNA	c 912	16	12.1	110000	2	AC026388	1
840	16	12.1	86696	2	AC097769	Rattus no	c 913	16	12.1	110000	2	AC026388	2
841	16	12.1	86774	9	AL359766	Human DNA	c 914	16	12.1	110000	2	AC027647	1
842	16	12.1	87286	2	AC014436	Drosophil	c 915	16	12.1	110157	8	AF058825	Arabidops
843	16	12.1	87402	9	HS3437M21	Human DNA	c 916	16	12.1	110630	2	OSJN00140	
844	16	12.1	87524	2	AC097876	Rattus no	c 917	16	12.1	110970	9	AF165144	Homo sapi
845	16	12.1	87822	2	AC013928	Drosophil	c 918	16	12.1	111081	2	AC010249	Homo sapi
846	16	12.1	88204	2	AC127210	Rattus no	c 919	16	12.1	111688	2	AC105020	Homo sapi
847	16	12.1	88278	2	AC015486	Homo sapi	c 920	16	12.1	111864	9	AL591499	Human DNA
848	16	12.1	89355	2	AC099423	Rattus no	c 921	16	12.1	112099	9	AL162727	Human DNA
849	16	12.1	89147	2	AC107911	Homo sapi	c 922	16	12.1	112365	9	AC027317	Homo sapi
850	16	12.1	89728	5	AL672079	Zebrafish	c 923	16	12.1	113172	8	AC009519	Genomic s
851	16	12.1	91433	3	AC087074	Caenorhab	c 924	16	12.1	113254	9	AL137009	Human DNA
852	16	12.1	92001	9	AC117476	Homo sapi	c 925	16	12.1	113601	2	AP003887	Oryza sat
853	16	12.1	92369	2	AC128342	Rattus no	c 926	16	12.1	114023	2	HS316D5	282199
854	16	12.1	92444	2	AC111809	Rattus no	c 927	16	12.1	114100	9	AL583806	Human DNA
855	16	12.1	92444	2	AC111809	Rattus no	c 928	16	12.1	114411	9	AC112909	Homo sapi
856	16	12.1	93260	2	AC127067	Rattus no	c 929	16	12.1	114597	2	AC129877	Rattus no
857	16	12.1	94023	9	AC008720	Homo sapi	c 930	16	12.1	115499	2	AC115363	Human DNA
858	16	12.1	94652	2	AC108876	Oryza sat	c 931	16	12.1	115614	9	AL139158	Human DNA
859	16	12.1	95105	9	AC016561	Homo sapi	c 932	16	12.1	116184	2	AC121241	Medicago
860	16	12.1	95660	2	AC127620	Rattus no	c 933	16	12.1	116380	9	AL1031671	Human DNA
861	16	12.1	95908	9	AC008628	Homo sapi	c 934	16	12.1	117029	2	AC101119	Rattus no
862	16	12.1	96157	9	AL138789	Human DNA	c 935	16	12.1	118129	9	AC093601	Homo sapi
863	16	12.1	96232	2	AP001921	Homo sapi	c 936	16	12.1	118334	9	AC005075	Homo sapi
864	16	12.1	96717	9	AL359182	Human DNA	c 937	16	12.1	118841	9	AC107973	Homo sapi
865	16	12.1	96779	9	AL592463	Human DNA	c 938	16	12.1	118988	2	AC109040	Rattus no
866	16	12.1	97059	2	AC129683	Rattus no	c 939	16	12.1	119395	9	AC091862	Homo sapi
867	16	12.1	97137	2	AL355172	Homo sapi	c 940	16	12.1	120228	2	AC112514	Oryza sat
868	16	12.1	97598	2	AC108238	Rattus no	c 941	16	12.1	120391	8	AC017146	Drosophil

942	16	12.1	120586	9	AL590322	Human DNA
c 943	16	12.1	120733	9	AC022124	Homo sapi
944	16	12.1	121360	9	AL157359	Homo sapi
945	16	12.1	121591	2	AC103260	Rattus no
c 946	16	12.1	121931	9	AC008390	Homo sapi
c 947	16	12.1	122228	9	AC000003	Homo sapi
c 948	16	12.1	122335	2	AC120636	Rattus no
c 949	16	12.1	122823	9	AC005996	Homo sapi
c 950	16	12.1	122884	9	HS1158B12	Human DNA
c 951	16	12.1	123022	9	AC090717	Homo sapi
c 952	16	12.1	123554	9	AB023049	Homo sapi
c 953	16	12.1	123668	9	AL390728	Human DNA
c 954	16	12.1	123925	9	AC003976	Homo sapi
c 955	16	12.1	124067	9	AP000435	Homo sapi
c 956	16	12.1	125267	4	AC090033	Felis cat
c 957	16	12.1	125528	2	AC027389	Homo sapi
c 958	16	12.1	125853	2	AP000613	Homo sapi
c 959	16	12.1	126130	9	AP000613	Homo sapi
c 960	16	12.1	126837	9	AL353652	Human DNA
c 961	16	12.1	126921	2	AC130576	Rattus no
c 962	16	12.1	126936	2	AC090472	Homo sapi
c 963	16	12.1	126990	9	HS73M23	Human DNA
c 964	16	12.1	127016	2	AC090730	Homo sapi
c 965	16	12.1	127119	2	AC123277	Rattus no
c 966	16	12.1	127144	9	AC008883	Homo sapi
c 967	16	12.1	127905	9	AL513550	Human DNA
c 968	16	12.1	128036	2	AC025071	Homo sapi
c 969	16	12.1	128227	2	AC118189	Rattus no
c 970	16	12.1	128364	2	AC122094	Rattus no
c 971	16	12.1	128437	9	AC084346	Homo sapi
c 972	16	12.1	128879	8	AP003866	Oryza sat
c 973	16	12.1	129096	2	AC026412	Homo sapi
c 974	16	12.1	129655	2	AC109027	Rattus no
c 975	16	12.1	130278	9	AC005664	Homo sapi
c 976	16	12.1	130336	9	AC006346	Homo sapi
c 977	16	12.1	130434	2	AC118459	Homo sapi
c 978	16	12.1	130561	9	AC107983	Homo sapi
c 979	16	12.1	130609	2	AC091798	Felis cat
c 980	16	12.1	130871	2	AC116517	Mus muscu
c 981	16	12.1	130898	2	AC130057	Rattus no
c 982	16	12.1	130898	2	AC130057	Rattus no
c 983	16	12.1	130935	2	AP004192	Oryza sat
c 984	16	12.1	131033	5	AL591370	Human DNA
c 985	16	12.1	131547	2	AP003982	Oryza sat
c 986	16	12.1	131805	9	AC008850	Homo sapi
c 987	16	12.1	132075	2	AC112539	Rattus no
c 988	16	12.1	132336	2	AF215842	Homo sapi
c 989	16	12.1	132367	10	AL808022	Mouse DNA
c 990	16	12.1	132933	9	AL137002	Human DNA
c 991	16	12.1	133927	9	AL512664	Human DNA
c 992	16	12.1	134933	8	AC092750	Oryza sat
c 993	16	12.1	134976	8	AP004968	Lotus jap
c 994	16	12.1	135237	2	AP000814	Homo sapi
c 995	16	12.1	135250	2	AC097048	Rattus no
c 996	16	12.1	135331	9	AC008798	Homo sapi
c 997	16	12.1	135360	2	AC126311	Rattus no
c 998	16	12.1	135815	9	HS43408	Human DNA
c 999	16	12.1	135888	2	AL845353	Homo sapi
c1000	16	12.1	135889	2	AP005469	Oryza sat
RESULT 1						
LOCUS	AX399230			147 bp		
DEFINITION	Sequence 3 from Patent WO0190155.					
ACCESSION	AX399230					
VERSION	AX399230.1			GI:21261549		
KEYWORDS	human.					
SOURCE	Homo sapiens					
ORGANISM						
1						
2						
3						
4						
5						
6						
7						
8						
9						
10						
11						
12						
13						
14						
15						
16						
17						
18						
19						
20						
21						
22						
23						
24						
25						
26						
27						
28						
29						
30						
31						
32						
33						
34						
35						
36						
37						
38						
39						
40						
41						
42						
43						
44						
45						
46						
47						
48						
49						
50						
51						
52						
53						
54						
55						
56						
57						
58						
59						
60						
61						
62						
63						
64						
65						
66						
67						
68						
69						
70						
71						
72						
73						
74						
75						
76						
77						
78						
79						
80						
81						
82						
83						
84						
85						
86						
87						
88						
89						
90						
91						
92						
93						
94						
95						
96						
97						
98						
99						
100						
101						
102						
103						
104						
105						
106						
107						
108						
109						
110						
111						
112						
113						
114						
115						
116						
117						
118						
119						
120						
121						
122						
123						
124						
125						
126						
127						
128						
129						
130						
131						
132						
133						
134						
135						
136						
137						
138						
139						
140						
141						
142						
143						
144						
145						
146						
147						
148						
149						
150						
151						
152						
153						
154						
155						
156						
157						
158						
159						
160						
161						
162						
163						
164						
165						
166						
167						
168						
169						
170						
171						
172						
173						
174						
175						
176						
177						
178						
179						
180						
181						
182						
183						
184						
185						
186						
187						
188						
189						
190						
191						
192						
193						
194						
195						
196						
197						
198						
199						
200						
201						
202						
203						
204						
205						
206						
207						
208						
209						
210						
211						
212						
213						
214						
215						
216						
217						
218						
219						
220						
221						
222						
223						
224						
225						
226						
227						

```

RESULT 3
AX399234
LOCUS       AX399234               147 bp    DNA        linear    PAT 03-JUN-2002
DEFINITION   Sequence 7 from Patent WO0190155.
ACCESSION   AX399234
VERSION     AX399234.1   GI:21261551
KEYWORDS
SOURCE      human.
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE   1
AUTHORS     Spytke,K.A., Majumder,K., Tchernev,V.T., Mishra,V., Padigaru,M.,
            Spaderna,S.K., Shenoy,S., Rastelli,L., Li,L., Taupier,R.J. and
            Gangolli,E.
TITLE       Novel proteins and nucleic acids encoding same
JOURNAL     Patent: WO 0190153-A 7 29-NOV-2001;
            Curagen Corporation (US)
FEATURES
            Location/Qualifiers
            1..147
                /organism="Homo sapiens"
                /db_xref="taxon:9606"
BASE COUNT   56 a   30 c   41 g   20 t
ORIGIN
Query Match      100.0%; Score 132; DB 6; Length 147;
Best Local Similarity 100.0%; Pred. No. 1e-66;
Matches 132; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGGCACAACAACTAGACCTGGGAAGAAATTCGACGTTGGATAGGCCAAGCTGAAGGCC 60
Db 1 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 6 ATGGCACAACAACTAGACCTGGGAAGAAATTCGACGTTGGATAGGCCAAGCTGAAGGCC 65
Db 6 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 61 ACAGAGATCGAGAGAACACTCTGATGACCAAGAACACACAGACGAGGAGAGTGGAGT 120
Db 61 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 66 ACAGAGATCGAGAGAACACTCTGATGACCAAGAACACACAGACGAGGAGAGTGGAGT 125
Db 66 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 121 GAAATTTCTCTGA 132
Db 121 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 126 GAAATTTCTCTGA 137
Db 126 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 4
HS518K17
LOCUS       HS518K17               170536 bp    DNA        linear    PRI 04-FEB-2002
DEFINITION   Homo sapiens chromosome 9 BAC RP11-518K17, complete sequence.
ACCESSION   AL513423
VERSION     AL513423.2   GI:18539155
KEYWORDS    HTG.
SOURCE      human.
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE   1
AUTHORS     Plumb,B.
TITLE       Direct Submission
JOURNAL     Submitted (09-SEP-2000) Sanger Centre, Hinxton, Cambridgeshire,
            CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
            Requests: clonerequest@sanger.ac.uk
            2 (bases 1 to 170536)
            Schaefer,M., Conrad,A., Hornischer,K., Loehnert,T.H., Thies,S. and
            Bloecker,H.
            Direct Submission
            Submitted (07-FEB-2001) GBF, Dept. of Genome Analysis, Mascheroder
            Weg 1, D-38124 Braunschweig, Germany, E-mail: info.genome@gbf.de
            On Feb 5, 2002 this sequence version replaced gi:12718191.
            All annotations in this database entry are developed by
            computational tools. It is therefore not explicitly noted in the
            feature lines that evidence is not experimental.
            Mapping was performed at The Sanger Centre
            (cf. http://www.sanger.ac.uk/HGP/Chr9)

```

Mapping information is available via
<http://webace.sanger.ac.uk/cgi-bin/display?db=acedb&gprep=518k17>

```

----- Genome Center
Center: GBF, Braunschweig
Center code: GBF
Web site: http://genome.gbf.de/
Contact: info.genome@gbf.de
----- Project Information
Center project name:
Center clone name: BA518K17
----- Summary Statistics
Sequencing vector: ###;
Chemistry: Dye-terminator-BigDye: 33% of reads
Chemistry: Dye-terminator-amersham: 55% of reads
Chemistry: Dye-primer-amersham: 12% of reads
Assembly program: Phrap: version 0.950319
Consensus quality: 0 bases at least Q40
Consensus quality: 0 bases at least Q30
Consensus quality: 0 bases at least Q20
Estimated insert size: ##; agarose-gel estimation
Estimated insert size: 170536; sum-of-contigs estimation
-----
PROGRAMS AND PARAMETERS USED FOR ANNOTATION:
+++++
+ Analysis and annotation were performed with the automatic
+ 'first-pass' annotation and submission tool
+ 'AnnoMitter' (Hornischer & Bloecker).
+ Programs used by 'AnnoMitter':
+ +++++
+ GeneFinder (Green), Vers. 084
+ Organism: human
+ GenScan (Burge & Karlin), Vers. 1.0
+ Used matrix: vertebrate; Minimum score: 0
+ Grail (Xu et al.), Vers. 1.3
+ Organism: human
+ Mzef (Zhang)
+ Prior probability: 0.04; Overlapping number: 0 > Xpound (Thomas
  & Skolnick)
BLASTN 2.0.14 (Altschul et al.)
+ Database(s): * RepBase: ALU (human), released 22-DEC-1995
+ * RepBase: LI (primate), released 22-DEC-1995
+ RepBase: MIR (primate), released 22-DEC-1995
+ RepBase: MER (primate), released 22-DEC-1995
+ RepBase: THE (primate), released 22-DEC-1995
+ Minimum identity: 70 %; Minimum score: 60;
+ * ESTs: BLASTN 2.0.14 (Altschul et al.)
+ Database(s): * emb1 (EST, human), released -DEC-
+ * emb1 (EST, other), released -DEC-
+ (EST), Vers. 67+ (01-JAN-1970). Using sequence with masked
  repeats
+ Minimum score: 60; Minimum identity: 70 %;
+ * Tandem Repeats: GDE 2.2 option 'tandem'
+ Minimum length 2 bp; Maximum length 20 bp; Score threshold 20
+ Treat N's as mismatches? YES; Allow uniform consensi? NO >
+ Inverted Repeats: GDE 2.2 option 'inverted'
+ * Micro Satellites: GDE 2.2 option 'sputnik' (Abaajian) > 'CpG
  Islands'. GDE 2.2 option 'cpg'
+ CpG island region size 100 bp;
+ Minimum GC contents 50 %; Observed/Expected 0.6 > 'STS Scan';
  e-PCR (Schuler)
+ Margin: 50; Number of mismatches allowed: 0; Word size: 7
+ STS database: 'dbSTS markers'
+ * 'TRNA Scan': tRNAscan-SE (Lowe & Eddy), Vers. 1.11.
  Location/Qualifiers
  1..170536
      /organism="Homo sapiens"
      /db_xref="taxon:9606"
      /chromosome="9"
      /clone_lib="RP11-518K17"
      /clone_lib="RPC1-11.2"

```

FEATURES
 source

Sequencing vector: M13; M7815; 9% of reads
 Sequencing vector: plasmid; L08752; 90% of reads
 Chemistry: Dye-terminator; 12% of reads
 Chemistry: Dye-terminator ABI; 0% of reads
 Chemistry: Dye-terminator ET-amersham; 0% of reads
 Dye-terminator Big Dye; 86% of reads
 Consensus quality: 86046 bases at least Q40
 Consensus quality: 86121 bases at least Q30
 Consensus quality: 86246 bases at least Q20
 Insert size: 86494; sum-of-contigs
 Insert size: 100607; 2.1% error; agarose-fp
 Quality coverage: 10.68x in Q20 bases; sum-of-contigs Quality
 coverage: 9.68x in Q20 bases; agarose-fp

 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 2 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

 * 1 46382: contig of 46382 bp in length
 * 46383 46482: gap of 100 bp
 * 46483 86594: contig of 40112 bp in length.

FEATURES

source

1. .86594
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="1"
 /clone="RP5-832C2"
 /clone_lib="RPCI-5"

misc_feature

1. .46382
 /note="assembly fragment:01669
 fragment_chain:1"

misc_feature

46483..86594
 /note="assembly fragment:00659
 fragment_chain:1
 clone_end:SP6
 vector_side:right"

BASE COUNT 19164 a 24662 c 24281 g 18387 t 100 others

ORIGIN

Query Match 18.9%; Score 25; DB 2; Length 86594;

Best Local Similarity 100.0%; Pred. No. 0.0021;

Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 106 CAGGAGAGTGGAGTGAATTCCT 130

|||||

Db 76479 CAGGAGAGTGGAGTGAATTCCT 76503

RESULT 6

AL713970

LOCUS AL713970 133135 bp DNA linear HTG 19-JUN-2002

DEFINITION Homo sapiens chromosome 1 clone RP11-197P23, *** SEQUENCING IN

PROGRESS ***, 15 unordered pieces.

ACCESSION AL713970

VERSION AL713970

KEYWORDS HTG; HTGS_PHASE1; HTGS_CANCELLED.

SOURCE human.

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 133135)

Ellington, A.

Direct Submission

Submitted (17-JUN-2002) Wellcome Trust Sanger Institute, Hinxton,

Cambridgeshire, CB10 1SA, UK. E-mail enquiries:

humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk

On Jun 21, 2002 this sequence version replaced gi:20451552.

----- Genome Center

Center: Wellcome Trust Sanger Institute

Center code: SC
 Web site: http://www.sanger.ac.uk
 Contact: humquery@sanger.ac.uk

----- Project Information

Center project name: BA197P23

----- Summary Statistics

Assembly program: XGAP4; version 4.5

Chemistry: Dye-terminator Big Dye; 100% of reads

Consensus quality: 127939 bases at least Q40

Consensus quality: 128929 bases at least Q30

Consensus quality: 129574 bases at least Q20

Insert size: 131735; sum-of-contigs

Insert size: 169742; 13.7% error; agarose-fp

Quality coverage: 7.78x in Q20 bases; sum-of-contigs Quality

coverage: 6.17x in Q20 bases; agarose-fp

* NOTE: This is a 'working draft' sequence. It currently
 * consists of 15 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

 * 1 4994: contig of 4994 bp in length
 * 4995 5094: gap of 100 bp
 * 5095 25519: contig of 20425 bp in length
 * 25520 25619: gap of 100 bp
 * 25620 35375: contig of 9756 bp in length
 * 35376 35475: gap of 100 bp
 * 35476 37722: contig of 2247 bp in length
 * 37723 37822: gap of 100 bp
 * 37823 49560: contig of 11738 bp in length
 * 49561 49660: gap of 100 bp
 * 49661 53726: contig of 4066 bp in length
 * 53727 53826: gap of 100 bp
 * 53827 84978: contig of 31152 bp in length
 * 84979 85078: gap of 100 bp
 * 85079 87297: contig of 2219 bp in length
 * 87298 87397: gap of 100 bp
 * 87398 100683: contig of 13286 bp in length
 * 100684 100783: gap of 100 bp
 * 100784 104917: contig of 4134 bp in length
 * 104918 105017: gap of 100 bp
 * 105018 114222: contig of 9205 bp in length
 * 114223 114322: gap of 100 bp
 * 114323 119300: contig of 4978 bp in length
 * 119301 119400: gap of 100 bp
 * 119401 122573: contig of 3173 bp in length
 * 122574 122673: gap of 100 bp
 * 122674 130806: contig of 8133 bp in length
 * 130807 130906: gap of 100 bp
 * 130907 133135: contig of 2229 bp in length.

FEATURES

source

1. .133135
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="1"
 /clone="RP11-197P23"
 /clone_lib="RPCI-11.1"

misc_feature

1. .4994
 /note="assembly fragment:02932
 fragment_chain:1"

misc_feature

5095..25519
 /note="assembly fragment:01032
 fragment_chain:1"

misc_feature

25620..35375
 /note="assembly fragment:01141
 fragment_chain:1"

misc_feature

35476..37722
 /note="assembly fragment:01459
 fragment_chain:1"

misc_feature

37823..49560


```

/note="assembly_fragment:00833
fragment_chain:1"
misc_feature
/note="assembly_fragment:00092
fragment_chain:2"
misc_feature
53827..84978
/note="assembly_fragment:02950
fragment_chain:2"
misc_feature
85079..87297
/note="assembly_fragment:00125
fragment_chain:2"
misc_feature
87398..100683
/note="assembly_fragment:01435
fragment_chain:2"
misc_feature
100784..104917
/note="assembly_fragment:01582
fragment_chain:3"
misc_feature
105018..114222
/note="assembly_fragment:01477
fragment_chain:3"
misc_feature
114323..119300
/note="assembly_fragment:01864
fragment_chain:3"
misc_feature
119401..122573
/note="assembly_fragment:00334
fragment_chain:4"
misc_feature
122674..130806
/note="assembly_fragment:00597
fragment_chain:4"
misc_feature
130907..133135
/note="assembly_fragment:01905"
BASE COUNT 29794 a 36915 c 36289 g 28728 t 1409 others
ORIGIN

```

Query Match 18.9%; Score 25; DB 2; Length 133135;

Best Local Similarity 100.0%; Pred. No. 0.0022; Mismatches 0; Indels 0; Gaps 0;

```

QY 106 CAGGAGAAGTGGAGTGAATTTCT 130
|||||
Db 82938 CAGGAGAAGTGGAGTGAATTTCT 82962

```

```

RESULT 7
AL691446 151840 bp DNA linear HTG 24-JUL-2002
LOCUS
DEFINITION
MUS musculus chromosome 2 clone RP23-136J11, *** SEQUENCING IN
PROGRESS ***, 10 unordered pieces.
ACCESSION AL691446
VERSION AL691446.2 GI:21531337
KEYWORDS HTG; HTGS_PHASE1; HTGS_ACTIVEFIN; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE Mus musculus.
ORGANISM Mus musculus.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
Bates, K.
Direct Submission
Submitted (23-JUL-2002) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
humquies@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
On Jun 21, 2002 this sequence version replaced gi:19572633.
----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: humquies@sanger.ac.uk
----- Project Information
Center project name: bml136J11
----- Summary Statistics
Assembly program: XGAP4; version 4.5
Chemistry: Dye-terminator; 0% of reads
Chemistry: Dye-terminator B1g Dye; 99% of reads

```

Consensus quality: 148339 bases at least Q40
 Consensus quality: 149486 bases at least Q30
 Consensus quality: 150246 bases at least Q20
 Insert size: 150940; sum-of-contigs
 Quality coverage: 11.92x in Q20 bases; sum-of-contigs Quality
 coverage: 7.78x in Q20 bases; agarose-fp

* NOTE: This is a 'working draft' sequence. It currently
 * consists of 10 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

```

* 1
* 2779 2878: gap of 100 bp
* 2879 8919: contig of 6041 bp in length
* 8920 9019: gap of 100 bp
* 9020 18791: contig of 9772 bp in length
* 18792 18891: gap of 100 bp
* 18892 30426: contig of 11535 bp in length
* 30427 30526: gap of 100 bp
* 30527 64454: contig of 33928 bp in length
* 64455 64554: gap of 100 bp
* 64555 85715: contig of 21161 bp in length
* 85716 85815: gap of 100 bp
* 85816 89857: contig of 4042 bp in length
* 89858 89957: gap of 100 bp
* 89958 106785: contig of 16828 bp in length
* 106786 106885: gap of 100 bp
* 106886 113841: contig of 6956 bp in length
* 113842 113941: gap of 100 bp
* 113942 151840: contig of 37899 bp in length.

```

FEATURES
 source

```

misc_feature
1..2778
/note="assembly_fragment:02749
fragment_chain:1"
misc_feature
2879..8919
/note="assembly_fragment:01224
fragment_chain:1"
misc_feature
9020..18791
/note="assembly_fragment:03159
fragment_chain:1"
misc_feature
18892..30426
/note="assembly_fragment:03065
fragment_chain:1"
misc_feature
30527..64454
/note="assembly_fragment:01666
fragment_chain:2"
misc_feature
64555..85715
/note="assembly_fragment:04137
fragment_chain:2"
misc_feature
85816..89857
/note="assembly_fragment:02847
fragment_chain:3"
misc_feature
89958..106785
/note="assembly_fragment:03601
fragment_chain:3"
misc_feature
106886..113841
/note="assembly_fragment:05543
fragment_chain:4"
misc_feature
113942..151840
/note="assembly_fragment:04887
fragment_chain:4"
BASE COUNT 34703 a 42897 c 43283 g 30057 t 900 others

```


/note="assembly_name:Contig61"

```

BASE COUNT 37019 a 48918 c 48934 g 40853 t 2008 others
ORIGIN
Query Match 18.9%; Score 25; DB 2; Length 177632;
Best Local Similarity 100.0%; Pred. No. 0.0022;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 106 CAGGAGAGTGGAGTGAATTCCT 130
|||||
Db 150599 CAGGAGAGTGGAGTGAATTCCT 150575

RESULT 9
AC097209
LOCUS Rattus norvegicus clone CH230-209D24, *** SEQUENCING IN PROGRESS
DEFINITION *** 36 unordered pieces.
ACCESSION AC097209.4 GI:21735659
VERSION HTG; HTGS_PHASE1.
KEYWORDS Norway rat.
SOURCE Rattus norvegicus
ORGANISM Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.

REFERENCE
1 (bases 1 to 60372)
Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-Osman,F.R., Allen,C.,
Albrooks,S.L., Amaraturge,H.C., Are,J.R., Ayele,M., Banks,T.,
Barbala,J., Benton,J., Binage,K., Blankenburg,K., Bonnin,D.,
Bouck,J., Bowle,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P.,
Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,
Carron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,
Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,
Cleveland,C.D., Cox,C., Coyne,M.D., Dathorne,S.R., David,R.,
Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,
Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,
Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,
Earhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,
Falls,T.; Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,
Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,
Gorell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,
Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J.,
Hernandez,O., Hodgson,A., Hognes,M., Holloway,C., Hollins,B.,
Honsi,F., Howard,S., Huber,J., Hulyk,S., Hume,J., Jackson,L.E.,
Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S.,
Karlisson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C.,
Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.,
Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W., Louseged,H.,
Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,
Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E.,
Massey,E., Mawhinney,E., McLeod,M.P., Meador,M., Mei,G., Metzker,M.,
Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S.,
Moser,M., Neal,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N.,
Nguyen,N., Nickerson,E., Nwokenwo,S., Oguh,M., Okwuonu,G.,
Oragunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L.,
Peters,L., Pickens,R., Primus,E., Pu,J.L., Quiles,M., Ren,Y.,
Rivers,M., Rojas,A., Rojibokan,I., Rolfe,M., Ruiz,S., Savery,G.,
Scherer,S., Scott,G., Shen,H., Shoohtari,N., Sisson,I.,
Sodergren,E., Sonaike,T., Sparks,A., Stanley,H., Stone,H.,
Sutton,A., Svatek,A., Taber,P., Tamerisa,A., Tamerisa,K., Tang,H.,
Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S.,
Usmani,K., Vasquez,L., Vera,V., Villalon,D., Vinson,R., Wang,Q.,
Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S.,
Williams,G., Williamson,A., Wleczyk,R., Wooden,S., Worley,K.,
Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,
Weinstock,G. and Gibbs,R.
Direct Submission
Unpublished
2 (bases 1 to 60372)
Worley,K.C.
Direct Submission
Submitted (12-OCT-2001) Human Genome Sequencing Center, Department

```

of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 60372)
Worley,K.C.
Direct Submission
Submitted (13-JUL-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Jul 12, 2002 this sequence version replaced gi:17973805.

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu

Project Information
Center project name: GIMK
Center clone name: CH230-209D24

Summary Statistics
Sequencing vector: Plasmid;

Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 18096 bases at least Q40
Consensus quality: 20317 bases at least Q30
Consensus quality: 22036 bases at least Q20

* NOTE: Estimated insert size may differ from sequence length
(see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 36 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 1116: contig of 1116 bp in length
1117 1216: gap of unknown length
1217 2301: contig of 1085 bp in length
2302 2401: gap of unknown length
2402 3901: contig of 1500 bp in length
3902 4001: gap of unknown length
4002 5011: contig of 1010 bp in length
5012 5112: gap of unknown length
5112 6141: contig of 1029 bp in length
6141 7365: gap of unknown length
7365 7465: contig of 1125 bp in length
7466 8533: contig of 1068 bp in length
8534 8634: gap of unknown length
8634 9937: contig of 1303 bp in length
9937 10036: gap of unknown length
10037 11520: contig of 1484 bp in length
11521 11621: gap of unknown length
11621 12664: contig of 1043 bp in length
12664 12764: gap of unknown length
12764 13845: contig of 1081 bp in length
13845 13945: gap of unknown length
13945 15744: contig of 1800 bp in length
15745 15845: gap of unknown length
15845 17261: contig of 1416 bp in length
17261 17361: gap of unknown length
17361 18647: contig of 1286 bp in length
18647 20062: contig of 1316 bp in length
20062 20163: gap of unknown length
20163 21879: contig of 1717 bp in length
21880 21980: gap of unknown length
21980 23605: contig of 1625 bp in length
23605 25341: gap of unknown length
25341 25441: contig of 1636 bp in length
25441 27219: gap of unknown length
27219 27318: contig of 1778 bp in length
27318: gap of unknown length

```

* 27319 28832: contig of 1514 bp in length
* 28832: gap of unknown length
* 28933: contig of 1598 bp in length
* 30530: contig of 1598 bp in length
* 30531: gap of unknown length
* 32775: contig of 2145 bp in length
* 32776: gap of unknown length
* 32876: contig of 1814 bp in length
* 34689: contig of 1814 bp in length
* 34690: gap of unknown length
* 36643: contig of 1854 bp in length
* 36644: gap of unknown length
* 39177: contig of 2434 bp in length
* 39178: gap of unknown length
* 39278: contig of 1357 bp in length
* 40634: gap of unknown length
* 40635: contig of 1879 bp in length
* 42613: gap of unknown length
* 42614: contig of 2511 bp in length
* 45224: gap of unknown length
* 45225: contig of 2336 bp in length
* 47660: gap of unknown length
* 47661: contig of 2035 bp in length
* 49795: gap of unknown length
* 49796: contig of 2222 bp in length
* 49896: gap of unknown length
* 52117: contig of 1613 bp in length
* 52118: gap of unknown length
* 52217: contig of 1576 bp in length
* 53830: contig of 1576 bp in length
* 53831: gap of unknown length
* 55507: contig of 2087 bp in length
* 55607: gap of unknown length
* 57694: contig of 1019 bp in length
* 57794: gap of unknown length
* 58812: contig of 1460 bp in length
* 58913: 60372: contig of 1460 bp in length.
      Location/Qualifiers
        1. 60372
          /organism="Rattus norvegicus"
          /db_xref="taxon:10116"
          /clone="CH230-209D24"
BASE COUNT 16444 a 12089 c 11630 g 16679 t 3530 others
ORIGIN
FEATURES
  SOURCE
    Query Match 16.7%; Score 22; DB 2; Length 60372;
    Best Local Similarity 100.0%; Pred. No. 0.13;
    Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 93 AGAGACCACAGAGAGAGAGAG 114
   |||||
Db 23269 AGAGACCACAGAGAGAGAG 23290

RESULT 10
AC098254/c
LOCUS
DEFINITION
AC098254
ACCESSION
VERSION
AC098254.3 GI:21953975
KEYWORDS
SOURCE
ORGANISM
  Rattus norvegicus
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
  Rattus.
REFERENCE
  1 (bases 1 to 192314)
  Muzny D.M., Adams C., Adio-Oduola B., Ali-Osman, F.R., Allen C.,
  Alsbrooks S.L., Amaralunge H.C., Are J.R., Ayale M., Banks T.,
  Barbarta J., Benton J., Bimaga K., Blankenburg K., Bonnin D.,
  Bouck J., Bowie S., Brieva M., Brown E., Brown M., Bryant N.P.,
  Buha C., Burch P., Burkett C., Burrell K.L., Byrd N.C.,
  Caron T.F., Carter M., Cavazos S.R., Chacko J., Chavez D.,
  Chen G., Chen R., Chen Z., Chowdhry I., Christopoulos C.,
  Cleveland C.D., Cox C., Coyie M.D., Dathorne S.R., David R.,
  Davila M.L., Davis C., Davy-Carroll L., Dederich D.A.,

```

```

Delaney, K.R., Delgado, O., Denn, A.L., Ding, Y., Dinh, H.H.,
Douthwaite, K.J., Draper, H., Dugan-Rocha, S., Durbin, K.J.,
Earnhart, C., Edgar, D., Edwards, C.C., Elhaj, C., Escotto, M.,
Falls, T., Ferraguto, D., Flagg, N., Ford, J., Foster, P., Frantz, P.,
Gabisi, A., Gao, J., Garcia, A., Garner, T., Garza, N., Gill, R.,
Gorrell, J.H., Guevara, W., Gunaratne, P., Hale, S., Hamilton, K.,
Harris, C., Harris, K., Hart, M., Havlak, P., Hawes, A., Hernandez, J.,
Hernandez, O., Hodgson, A., Hogues, M., Holloway, C., Hollins, B.,
Homs, F., Howard, S., Huber, J., Hulyk, S., Hume, J., Jackson, L.E.,
Jacobson, B., Jia, Y., Johnson, R., Jolivet, S., Joudah, S.,
Karissom, E., Kelly, S., Khan, U., King, L., Korvan, J., Kovar, C.,
Kratovic, J., Kureshi, A., Landry, N., Leal, B., Lewis, L.C., Lewis, L.,
Li, J., Li, Z., Lichtarge, O., Lieu, C., Liu, J., Liu, W., Louisedge, H.,
Lozado, R.J., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J.,
Maheshwari, M., Mapua, P., Martin, R., Martindale, A., Martinez, E.,
Massey, E., Mawhiney, E., McLeod, M.P., Meador, M., Mei, G., Metzker, M.,
Miner, G., Miner, Z., Mitchell, F., Mohabbat, K., Morgan, M., Morris, S.,
Moser, M., Neal, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N.,
Nguyen, N., Nickerson, E., Nwokenkwo, S., Ogih, M., Okwuonu, G.,
Oragunye, N., Oviedo, R., Pace, A., Payton, B., Peery, J., Perez, L.,
Peters, L., Pickens, R., Primus, E., Pu, L.L., Quiles, M., Ren, Y.,
Rives, M., Rojas, A., Rojubokan, I., Rolfe, M., Ruiz, S., Savery, G.,
Scherer, S., Scott, G., Shen, H., Shoostari, N., Sisson, I.,
Sodergren, E., Sonaika, T., Sparks, A., Stanley, H., Stone, H.,
Sutton, A., Svatek, A., Tabor, P., Tamerisa, A., Tamerisa, K., Tang, H.,
Tansey, J., Taylor, C., Taylor, T., Telford, B., Thomas, N., Thomas, S.,
Usmani, K., Vasquez, L., Vera, V., Villalon, D., Vinson, R., Wang, Q.,
Wang, S., Ward-Moore, S., Warren, R., Washington, C., Watlington, S.,
Williams, G., Williamson, A., Wleczyk, R., Woodson, S., Worley, K.,
Wu, C., Wu, Y., Wu, Y.F., Zhou, J., Zorrilla, S., Nelson, D.,
Weinstock, G. and Gibbs, R.
Direct Submission
Unpublished
2 (bases 1 to 192314)
Worley, K.C.
Direct Submission
Submitted (23-OCT-2001) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 192314)
Worley, K.C.
Direct Submission
Submitted (31-JUL-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Jul 24, 2002 this sequence version replaced gi:20976509.
----- Genome Center of Medicine
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: TUDS
Center clone name: CH230-1G4
----- Summary Statistics
Sequencing vector: Plasmid;
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 124741 bases at least Q40
Consensus quality: 130791 bases at least Q30
Consensus quality: 137646 bases at least Q20
-----
* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 64 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1 1202: contig of 1202 bp in length

```

```

* 1203 1302: gap of unknown length
* 1303 2470: contig of 1168 bp in length
* 2471 2570: gap of unknown length
* 2571 3814: contig of 1244 bp in length
* 3815 3914: gap of unknown length
* 3915 5367: contig of 1453 bp in length
* 5368 5467: gap of unknown length
* 5468 6821: contig of 1354 bp in length
* 6822 6921: gap of unknown length
* 6922 8132: contig of 1211 bp in length
* 8133 8232: gap of unknown length
* 8233 9889: contig of 1657 bp in length
* 9890 11626: gap of unknown length
* 11627 11726: gap of unknown length
* 11727 13008: contig of 1282 bp in length
* 13009 13108: gap of unknown length
* 13109 14592: contig of 1484 bp in length
* 14593 14692: gap of unknown length
* 14693 15718: contig of 1026 bp in length
* 15719 15818: gap of unknown length
* 15819 17454: contig of 1636 bp in length
* 17455 17554: gap of unknown length
* 17555 19138: contig of 1584 bp in length
* 19139 19239: gap of unknown length
* 19239 20377: contig of 1138 bp in length
* 20377 20476: gap of unknown length
* 20477 21841: contig of 1365 bp in length
* 21842 21941: gap of unknown length
* 21942 23808: contig of 1867 bp in length
* 23809 23908: gap of unknown length
* 23909 25201: contig of 1292 bp in length
* 25201 25300: gap of unknown length
* 25301 26438: contig of 1138 bp in length
* 26439 26538: gap of unknown length
* 26539 27588: contig of 1050 bp in length
* 27589 27688: gap of unknown length
* 27689 29362: contig of 1673 bp in length
* 29362 30997: contig of 1536 bp in length
* 30998 31097: gap of unknown length
* 31098 32929: contig of 1832 bp in length
* 32930 33029: gap of unknown length
* 33030 34610: contig of 1581 bp in length
* 34611 34710: gap of unknown length
* 34711 36964: contig of 2254 bp in length
* 36965 37064: gap of unknown length
* 37065 38944: contig of 1880 bp in length
* 38945 39044: gap of unknown length
* 39045 40758: contig of 1714 bp in length
* 40759 40858: gap of unknown length
* 40859 42629: contig of 1771 bp in length
* 42630 42729: gap of unknown length
* 42730 44658: contig of 1929 bp in length
* 44659 44758: gap of unknown length
* 44759 46330: contig of 1572 bp in length
* 46331 48430: gap of unknown length
* 48431 48745: contig of 2315 bp in length
* 48746 48845: gap of unknown length
* 48846 50969: contig of 2124 bp in length
* 50970 51069: gap of unknown length
* 51070 52350: contig of 1281 bp in length
* 52351 52450: gap of unknown length
* 52451 54791: contig of 2341 bp in length
* 54792 54891: gap of unknown length
* 54892 56266: contig of 1375 bp in length
* 56267 56366: gap of unknown length
* 56367 58908: contig of 2542 bp in length
* 58909 59008: gap of unknown length
* 59009 61210: contig of 2202 bp in length
* 61211 61310: gap of unknown length
* 61311 64169: contig of 2859 bp in length
* 64170 64269: gap of unknown length

```

```

* 64270 65546: contig of 1277 bp in length
* 65547 65646: gap of unknown length
* 65647 67455: contig of 1809 bp in length
* 67456 67555: gap of unknown length
* 67556 71809: contig of 4254 bp in length
* 71810 71909: gap of unknown length
* 71910 74306: contig of 2397 bp in length
* 74307 74406: gap of unknown length
* 74407 76355: contig of 1949 bp in length
* 76356 76456: gap of unknown length
* 76457 79430: contig of 2975 bp in length
* 79431 79531: gap of unknown length
* 79532 83084: contig of 3553 bp in length
* 83085 83184: gap of unknown length
* 83185 87278: contig of 4095 bp in length
* 87279 87378: gap of unknown length
* 87379 89371: contig of 1993 bp in length
* 89372 89471: gap of unknown length
* 89472 93781: contig of 4310 bp in length
* 93782 93881: gap of unknown length
* 93882 96980: contig of 3099 bp in length
* 96981 97080: gap of unknown length
* 97081 100549: contig of 3469 bp in length
* 100550 102941: gap of unknown length
* 102942 103041: gap of unknown length
* 103042 107118: contig of 4077 bp in length
* 107119 107218: gap of unknown length
* 107219 110788: contig of 3570 bp in length
* 110789 110888: gap of unknown length
* 110890 115458: contig of 4570 bp in length
* 115459 115558: gap of unknown length
* 115559 119571: contig of 4013 bp in length

```

```

Query Match      16.7%: Score 22; DB 2; Length 192314;
Best Local Similarity 100.0%: Pred. No. 0.13;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 93 AGAGACCACAGCAGCAGGAGAAG 114
    |||||
Db 36812 AGAGACCACAGCAGCAGGAGAAG 36791

```

RESULT 11

```

AC130181
LOCUS      Rattus norvegicus clone CH230-5J1, linear HTG 08-AUG-2002
DEFINITION Rattus norvegicus clone CH230-5J1, *** SEQUENCING IN PROGRESS ***,
56 unordered pieces.
ACCESSION  AC130181
VERSION     AC130181.1 GI:22138425
KEYWORDS    HTG; HTGS_PHASE1.
SOURCE      Norway rat.
ORGANISM    Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.

```

REFERENCE

```

AUTHORS
1 (bases 1 to 198015)
Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,
Alsbrooks,S.L., Amaratunge,H.C., Are,J.R., Ayale,M., Banks,T.,
Barbaria,J., Benton,J., Bimage,K., Blankenburg,K., Bonnin,D.,
Bouck,J., Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P.,
Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,
Carron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chaves,D.,
Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,
Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,
Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,
Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,
Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,
Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,
Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,
Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,
Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,
Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J.,

```

Hernandez, O., Hodgson, A., Hognes, M., Holloway, C., Hollins, B.,
Honsi, P., Howard, S., Huber, J., Hulyk, S., Hume, J., Jackson, L. E.,
Jacobson, B., Jia, Y., Johnson, R., Jolivet, S., Joudah, S.,
Karlovic, E., Kelly, S., Khan, U., King, L., Korvah, J., Kovar, C.,
Kratovic, J., Kureshi, A., Landry, N., Leal, B., Lewis, L. C., Lewis, L.,
Li, J., Li, Z., Lichtarge, O., Lieu, C., Lucier, R., Luna, R., Ma, J.,
Lozano, R. J., Lu, X., Lucier, R., Lucier, R., Luna, R., Ma, J.,
Maheshwari, M., Mapua, P., Martin, R., Martindale, A., Martinez, E.,
Massey, L., Massey, E., McLeod, M. P., Meador, M., Mei, G., Metzker, M.,
Miner, G., Miner, Z., Mitchell, T., Mohabbat, K., Morgan, M., Morris, S.,
Moser, M., Neal, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N.,
Nguyen, N., Nickerson, E., Nwokenkwo, S., Ogih, M., Okwuonu, G.,
Orgunye, N., Oviedo, R., Pace, A., Payton, B., Peery, J., Perez, L.,
Peters, L., Pickens, R., Primus, E., Pu, L. L., Quiles, M., Ren, Y.,
Rives, M., Rojas, A., Rojibokan, I., Rolfe, M., Ruiz, S., Savary, G.,
Scherer, S., Scott, G., Shen, H., Shoostari, N., Sisson, I.,
Sodergren, E., Sonaike, T., Sparks, A., Stanley, H., Stone, H.,
Sutton, A., Svatek, A., Taboz, P., Tamerisa, A., Tang, H.,
Tansey, J., Taylor, C., Taylor, T., Teifrod, B., Thomas, N., Thomas, S.,
Usmani, K., Vasquez, L., Vera, V., Villalob, D., Vinson, R., Wang, Q.,
Wang, S., Ward-Moore, S., Warren, R., Washington, C., Watlington, S.,
Williams, G., Williamson, A., Wleczky, R., Woodden, S., Worley, K.,
Wu, C., Wu, Y., Wu, Y. F., Zhou, J., Zorrilla, S., Nelson, D.,
Weinstock, G., and Gibbs, R.

Direct Submission
Unpublished
2 (bases 1 to 198015)
Worley, K.C.
Direct Submission
Submitted (08-AUG-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA

Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu

Project Information
Center project name: GBLD
Center clone name: CH230-5J1

Summary Statistics
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 151044 bases at least Q40
Consensus quality: 160277 bases at least Q30
Consensus quality: 166259 bases at least Q20

* NOTE: Estimated insert size may differ from sequence length
(see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 56 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1536: contig of 1536 bp in length
1537: 1636: gap of unknown length
1637: 3117: contig of 1481 bp in length
3118: 3217: gap of unknown length
3218: 4484: contig of 1267 bp in length
4485: 4584: gap of unknown length
4585: 6182: contig of 1598 bp in length
6183: 6282: gap of unknown length
6283: 7357: contig of 1075 bp in length
7358: 7457: gap of unknown length
7458: 9367: contig of 1910 bp in length
9368: 9467: gap of unknown length
9468: 10942: contig of 1475 bp in length
10943: 11042: gap of unknown length
11043: 13268: contig of 2226 bp in length
13269: 13368: gap of unknown length

13369: 15041: contig of 1673 bp in length
15042: 15141: gap of unknown length
15142: 16627: contig of 1486 bp in length
16628: 16727: gap of unknown length
16728: 17943: contig of 1216 bp in length
17944: 18043: gap of unknown length
18045: 19081: contig of 1038 bp in length
19082: 19181: gap of unknown length
19182: 20826: contig of 1545 bp in length
20827: 22566: gap of unknown length
22567: 22666: gap of unknown length
22668: 24418: contig of 1752 bp in length
24419: 24518: gap of unknown length
24519: 26895: contig of 2377 bp in length
26896: 26995: gap of unknown length
26996: 29066: contig of 2071 bp in length
29067: 29166: gap of unknown length
29167: 30375: contig of 1209 bp in length
30376: 30475: gap of unknown length
30476: 32675: contig of 2200 bp in length
32676: 32775: gap of unknown length
32776: 35034: contig of 2259 bp in length
35035: 35134: gap of unknown length
35135: 37314: contig of 2180 bp in length
37315: 37415: gap of unknown length
37416: 39405: contig of 1991 bp in length
39406: 42743: contig of 3238 bp in length
42744: 42843: gap of unknown length
42845: 45730: contig of 2887 bp in length
45731: 45830: gap of unknown length
45831: 47280: contig of 1450 bp in length
47281: 47380: gap of unknown length
47382: 50504: contig of 3124 bp in length
50505: 50604: gap of unknown length
50606: 53646: contig of 3042 bp in length
53647: 53746: gap of unknown length
53747: 56155: contig of 2409 bp in length
56156: 56255: gap of unknown length
56256: 57617: contig of 1362 bp in length
57618: 57717: gap of unknown length
57718: 59390: contig of 1673 bp in length
59391: 59490: gap of unknown length
59491: 61910: contig of 2420 bp in length
61911: 62010: gap of unknown length
62011: 64435: contig of 2425 bp in length
64436: 64535: gap of unknown length
64536: 66876: contig of 2341 bp in length
66877: 66976: gap of unknown length
66977: 70641: contig of 3665 bp in length
70642: 70741: gap of unknown length
70742: 74174: contig of 3433 bp in length
74175: 74274: gap of unknown length
74275: 77442: contig of 3168 bp in length
77443: 77542: gap of unknown length
77543: 79515: gap of unknown length
79516: 82758: contig of 3243 bp in length
82759: 82858: gap of unknown length
82859: 86464: contig of 3606 bp in length
86465: 86564: gap of unknown length
86565: 91223: contig of 4659 bp in length
91224: 91323: gap of unknown length
91324: 96124: contig of 4801 bp in length
96125: 96224: gap of unknown length
96225: 99485: contig of 3261 bp in length
99486: 99585: gap of unknown length
99586: 104436: contig of 4851 bp in length
104437: 104536: gap of unknown length
104537: 108432: contig of 3896 bp in length
108433: 108532: gap of unknown length
108533: 111718: contig of 3186 bp in length

TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

* 111719 111818: gap of unknown length
 * 111819 116364: contig of 4546 bp in length
 * 116365 116464: gap of unknown length
 * 116465 120991: contig of 4527 bp in length
 * 120992 121091: gap of unknown length
 * 121092 121178: contig of 6087 bp in length
 * 121179 127278: gap of unknown length
 * 127279 130963: contig of 3685 bp in length
 * 130964 131063: gap of unknown length
 * 131064 136356: contig of 5293 bp in length
 * 136357 136456: gap of unknown length
 * 136457 143165: contig of 6709 bp in length
 * 143166 143265: gap of unknown length
 * 143266 150981: contig of 7616 bp in length
 * 150982 163191: contig of 12210 bp in length
 * 163192 163291: gap of unknown length
 * 163292 172895: contig of 9604 bp in length
 * 172896 172995: gap of unknown length
 * 172996 185774: contig of 12779 bp in length
 * 185775 185874: gap of unknown length
 * 185875 198015: contig of 12141 bp in length.

FEATURES
 source
 1. 198015
 /organism="Rattus norvegicus"
 /db_xref="taxon:10116"

Query Match 15.9%; Score 21; DB 2; Length 198015;
 Best Local Similarity 100.0%; Pred. No. 0.51;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 93 AGAGCCACAGCAGGAGAA 113
 Db 42390 AGAGCCACAGCAGGAGAA 42410

RESULT 12
 AP001477 38399 bp DNA linear PRI 09-MAY-2000
 LOCUS
 DEFINITION Homo sapiens genomic DNA, chromosome 21q22.3, clone:f50C10,
 telomere region, complete sequence.
 ACCESSION AP001477
 VERSION AP001477.1 GI:7768605
 KEYWORDS HTG.
 SOURCE Homo sapiens DNA, clone:f50C10.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 38399)
 AUTHORS Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
 Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
 TITLE Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
 JOURNAL Published Only in DataBase (2000)
 REFERENCE 2 (bases 1 to 38399)
 AUTHORS Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
 Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
 TITLE Direct Submission
 JOURNAL Submitted (16-MAR-2000) Masahira Hattori, The Institute of Physical
 and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
 Kitasato Univ., 1-15-1 Kitasato, Sagamihara, Kanagawa 228-8555,
 Japan (E-mail:hattori@gs.c.riken.go.jp,
 URL:http://hgp.gsc.riken.go.jp/, Tel:81-42-778-9923,
 Fax:81-42-778-9924)

FEATURES
 source
 1. 38399
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="21"
 /map="21q22.3"
 /clone="f50C10"

BASE COUNT 7773 a 9530 c 10510 g 10586 t
 ORIGIN

Query Match 15.2%; Score 20; DB 9; Length 38399;
 Best Local Similarity 100.0%; Pred. No. 1.9;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 27 AATTGCCAGCTGGATAGG 46
 Db 26428 AATTGCCAGCTGGATAGG 26447

RESULT 13
 AC116359/c 53707 bp DNA linear PRI 29-MAY-2002
 LOCUS
 DEFINITION Homo sapiens chromosome 5 clone RP11-692C23, complete sequence.
 ACCESSION AC116359
 VERSION AC116359.2 GI:21240691
 KEYWORDS HTG.
 SOURCE Homo sapiens.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 53707)
 AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.
 TITLE Direct Submission
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 53707)
 AUTHORS DOE Joint Genome Institute.
 TITLE Direct Submission
 JOURNAL Submitted (27-MAR-2002) Production Sequencing Facility, DOE Joint
 Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
 REFERENCE 3 (bases 1 to 53707)
 AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.
 TITLE Direct Submission
 JOURNAL Submitted (29-MAY-2002) DOE Joint Genome Institute, 2800 Mitchell
 Drive, Walnut Creek, CA 94598, USA
 COMMENT On May 29, 2002 this sequence version replaced gi:19745041.
 Draft Sequence Produced by DOE Joint Genome Institute
 www.jgi.doe.gov

Finishing Completed at Stanford Human Genome Center
 www.shgc.stanford.edu
 Quality: Phrap Quality >=40 100% of Sequence;
 Estimated Total Number of Errors is 0.

NOTE: This insert is not the entire sequence of the clone (entire
 sequence is 171.7kb). It is clipped at the overlaps with AC094105
 and AC091891. the number of bases overlapped with AC094105 is 8849
 and with AC091891 is 44862.
 Location/Qualifiers
 1. 53707
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="5"
 /clone="RP11-692C23"

BASE COUNT 14758 a 10959 c 11116 g 16874 t
 ORIGIN

FEATURES
 source
 1. 53707
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="5"
 /clone="RP11-692C23"

BASE COUNT 14758 a 10959 c 11116 g 16874 t
 ORIGIN

Query Match 15.2%; Score 20; DB 9; Length 53707;
 Best Local Similarity 100.0%; Pred. No. 1.9;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 64 GAGATGCAGAGAACACTCT 83
 Db 49318 GAGATGCAGAGAACACTCT 49299

RESULT 14
 AC025937/c 170595 bp DNA linear HTG 26-MAY-2000
 LOCUS
 DEFINITION Homo sapiens clone RP11-692C23, WORKING DRAFT SEQUENCE, 18
 unordered pieces.
 ACCESSION AC025937
 VERSION AC025937.3 GI:8077762
 KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
 SOURCE Homo sapiens.
 ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 170595)
 Birren,B., Linton,L., Nusbaum,C. and Lander,E.
 Homo sapiens, clone RP11-692C23
 2 (bases 1 to 170595)
 Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N., Anderson,S., Baldwin,J., Barna,N., Bastien,V., Bada,F., Boguslavskiy,L., Boukhalter,B., Brown,A., Burkett,G., Campopiano,A., Castle,A., Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S., Domino,M., Doyle,M., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L., Grand-Pierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L., Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J., LaRoque,K., Lamazares,R., Landers,T., Lehoczy,J., Levine,R., Liu,G., Locke,K., Macdonald,P., Marquis,N., McCarthy,M., McEwan,P., McGurk,A., McKernan,K., McPheeters,R., Meldrim,J., Meneus,L., Mihova,T., Miranda,C., Mlenga,V., Morrow,J., Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,T.M., Oliver,J., Peterson,K., Pierre,N., Pisani,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D., Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J., Testaye,S., Theodore,J., Tirrell,A., Travers,M., Trigilio,J., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J., Zimmer,A. and Zody,M.
 Direct Submission
 Submitted (17-MAR-2000) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
 On May 25, 2000 this sequence version replaced gi:7408053.
 All repeats were identified using RepeatMasker:
 Smit, A.F.A. & Green, P. (1996-1997)
 http://ftp.genome.washington.edu/RM/RepeatMasker.html
 ----- Genome Center
 Center: Whitehead Institute/ MIT Center for Genome Research
 Center code: WIBR
 Web site: http://www-seq.wi.mit.edu
 Contact: sequence_submissions@genome.wi.mit.edu
 ----- Project Information
 Center project name: L8334
 Center clone name: 692_C_23
 ----- Summary Statistics
 Sequencing vector: M13; M77815; 100% of reads
 Chemistry: Dye-terminator Big Dye; 100% of reads
 Assembly program: Phrap; version 0.960731
 Consensus quality: 160465 bases at least Q40
 Consensus quality: 165270 bases at least Q30
 Consensus quality: 167427 bases at least Q20
 Insert size: 164000; agarose-fp
 Insert size: 168895; sum-of-contigs
 Quality coverage: 4.7 in Q20 bases; agarose-fp
 Quality coverage: 4.6 in Q20 bases; sum-of-contigs

 * NOTE: This is a 'working draft' sequence. It currently consists of 18 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.
 * 1 1144: contig of 1144 bp in length
 * 1145 1244: gap of 100 bp
 * 1245 2825: contig of 1581 bp in length
 * 2826 2925: gap of 100 bp
 * 2926 4246: contig of 1321 bp in length
 * 4247 4346: gap of 100 bp
 * 4347 6039: contig of 1693 bp in length
 * 6040 6139: gap of 100 bp
 * 6140 8144: contig of 2005 bp in length
 * 8145 8244: gap of 100 bp

* 8245 13130: contig of 4886 bp in length
 * 13131 13230: gap of 100 bp
 * 13231 16109: contig of 2879 bp in length
 * 16110 16209: gap of 100 bp
 * 16210 19026: contig of 2817 bp in length
 * 19027 19126: gap of 100 bp
 * 19127 24489: contig of 5363 bp in length
 * 24490 24589: gap of 100 bp
 * 24590 28306: contig of 3717 bp in length
 * 28307 28406: gap of 100 bp
 * 28407 38498: contig of 10092 bp in length
 * 38499 38598: gap of 100 bp
 * 38599 48302: contig of 9704 bp in length
 * 48303 48402: gap of 100 bp
 * 48403 57446: contig of 9044 bp in length
 * 57447 57546: gap of 100 bp
 * 57547 68273: contig of 10727 bp in length
 * 68274 68373: gap of 100 bp
 * 68374 79825: contig of 11452 bp in length
 * 79826 79925: gap of 100 bp
 * 79926 101549: contig of 21624 bp in length
 * 101550 101649: gap of 100 bp
 * 101650 130009: contig of 28360 bp in length
 * 130010 130109: gap of 100 bp
 * 130110 170595: contig of 40486 bp in length.
 FEATURES
 Location/Qualifiers
 source
 1..170595
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_lib="RPCI-11 Human Male BAC"
 misc_feature
 1..1144
 /note="assembly_fragment"
 misc_feature
 1245..2825
 /note="assembly_fragment"
 misc_feature
 2926..4246
 /note="assembly_fragment"
 misc_feature
 4347..6039
 /note="assembly_fragment"
 misc_feature
 6140..8144
 /note="assembly_fragment"
 misc_feature
 8245..13130
 /note="assembly_fragment"
 misc_feature
 13231..16109
 /note="assembly_fragment"
 misc_feature
 16210..19026
 /note="assembly_fragment"
 misc_feature
 19127..24489
 /note="assembly_fragment"
 misc_feature
 24590..28306
 /note="assembly_fragment"
 misc_feature
 clone_end:SP6
 vector_side:left
 28407..38498
 /note="assembly_fragment"
 misc_feature
 38599..48302
 /note="assembly_fragment"
 misc_feature
 48403..57446
 /note="assembly_fragment"
 misc_feature
 clone_end:17
 vector_side:right
 57547..68273
 /note="assembly_fragment"
 misc_feature
 68374..79825
 /note="assembly_fragment"
 misc_feature
 79926..101549
 /note="assembly_fragment"
 misc_feature
 101650..130009
 /note="assembly_fragment"
 misc_feature
 130110..170595
 /note="assembly_fragment"
 BASE COUNT 45763 a 37848 c 38643 g 46640 t 1701 others
 ORIGIN

Query Match 15.2%; Score 20; DB 2; Length 170595;
 Best Local Similarity 100.0%; Pred. No. 2;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 64 GAGATGCAGAGAACACTCT 83
 |||||
 Db 125491 GAGATGCAGAGAACACTCT 125472

RESULT 15
 AC113727/c
 LOCUS AC113727
 DEFINITION Rattus norvegicus clone CH230-4709, *** SEQUENCING IN PROGRESS ***,
 71 unordered pieces.
 AC113727
 VERSION AC113727.3 GI:21744507
 KEYWORDS HTG; HTGS-PHASE1.
 SOURCE Norway rat.
 ORGANISM Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;

REFERENCE
 1 (bases 1 to 187097)
 Muzny D.M., Adams C., Adio-Oduola B., Ali-Osman F.R., Allen C.,
 Albrooks S.L., Amarantunge H.C., Are J.R., Ayale M., Banks T.,
 Barbarella J., Benton J., Blmage K., Blankenburg K., Bonnin D.,
 Bouck J., Bowie S., Brileva M., Brown E., Brown M., Bryant N.P.,
 Buhay C., Burch P., Burkett C., Burrell K.L., Byrd N.C.,
 Carron T.F., Carter M., Cavazos S.R., Chacko J., Chavez D.,
 Chen G., Chen R., Chen Z., Chowdhry I., Christopoulos C.,
 Cleveland C.D., Cox C., Coyle M.D., Dathorne S.R., David R.,
 Davila M.L., Davis C., Davy-Carroll L., Dederich D.A.,
 Delaney K.R., Delgado O., Denn A.L., Ding Y., Dinh H.H.,
 Douthwaite K.J., Draper H., Dugan-Rocha S., Durbin K.J.,
 Earnhart C., Edgar D., Edwards C.C., Elhaj C., Escotto M.,
 Falls T., Ferraguto D., Flagg N., Ford J., Foster P., Frantz P.,
 Gabisi A., Gao J., Garcia A., Garner T., Garza N., Gill R.,
 Gorrill J.H., Guevara W., Gunaratne P., Hale S., Hamilton K.,
 Harris C., Harris K., Hart M., Havlak P., Hawes A., Hernandez J.,
 Hernandez O., Hodgson A., Hogues M., Holloway C., Hollins B.,
 Homs J., Howard S., Huber J., Hulyk S., Hume J., Jackson L.E.,
 Jacobson B., Jia Y., Johnson R., Jolivet S., Joudah S.,
 Karlsson E., Kelly S., Khan U., King L., Korvah J., Kovar C.,
 Kratovic J., Kuresh A., Landry N., Leal B., Lewis L.C., Lewis L.,
 Li J., Li Z., Lichtarge O., Lieu C., Lucier R., Luna R., Ma J.,
 Lozado R.J., Lu X., Lucier A., Lucier R., Martindale A., Martinez E.,
 Maheshwari M., Mapua P., Martin R., Mohabbat K., Morgan M., Morris S.,
 Massey E., Mawhiney E., McLeod M.P., Meador M., Mei G., Metzker M.,
 Miner G., Miner Z., Mitchell T., Mohabbat K., Morgan M., Morris S.,
 Moser M., Neal D., Newton J., Newton N., Nguyen A., Nguyen N.,
 Nguyen N., Nickerson E., Nwokenwo S., Oguh M., Okwuonu G.,
 Oragunye N., Oviedo R., Pace A., Payton B., Peery J., Perez L.,
 Peters L., Pickens R., Primus E., Pu L.L., Quiles M., Ren Y.,
 Rives M., Rojas A., Rojibokan I., Rolfe M., Ruiz S., Savery G.,
 Scherer S., Scott G., Shen H., Shooshtari N., Sisson I.,
 Sodergren E., Soneike T., Sparks A., Stanley H., Stone H.,
 Stott A., Svatek A., Tabor P., Tamerisa A., Tamerisa K., Tang H.,
 Tansey J., Taylor C., Taylor T., Telford B., Thomas N., Thomas S.,
 Usmani K., Vasquez L., Vera V., Villalón D., Vinson R., Wang Q.,
 Wang S., Ward-Moore S., Warren R., Washington C., Watlington S.,
 Williams G., Williamson A., Wleczyk R., Wooden S., Worley K.,
 Wu C., Wu Y., Wu Y.F., Zhou J., Zorrilla S., Nelson D.,
 Weinstein G., and Gibbs R.

TITLE Direct Submission
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 187097)
 AUTHORS Worley K.C.
 TITLE Direct Submission
 JOURNAL Submitted (05-MAR-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 REFERENCE 3 (bases 1 to 187097)

AUTHORS
 TITLE
 JOURNAL
 COMMENT

Worley, K.C.
 Direct Submission
 Submitted (17-JUL-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 On Jul 13, 2002 this sequence version replaced gi:19525893.
 ----- Genome Center
 Center: Baylor College of Medicine
 Center code: BCM
 Web site: <http://www.hgsc.bcm.tmc.edu/>
 Contact: hgsc-help@bcm.tmc.edu
 ----- Project Information
 Center project name: GSDG
 Center clone name: CH230-4709
 ----- Summary Statistics
 Sequencing vector: Plasmid;
 Chemistry: Dye-terminator Big Dye; 100% of reads
 Assembly program: Phrap; version 0.990329
 Consensus quality: 120207 bases at least Q40
 Consensus quality: 124896 bases at least Q30
 Consensus quality: 129455 bases at least Q20

* NOTE: Estimated insert size may differ from sequence length
 * (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 71 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

1 1170: contig of 1170 bp in length
 1171 1270: gap of unknown length
 1271 2275: contig of 1005 bp in length
 2276 3751: gap of unknown length
 3752 3851: contig of 1376 bp in length
 3852 4860: contig of unknown length
 4861 4960: gap of unknown length
 4961 6328: contig of 1368 bp in length
 6329 8022: gap of unknown length
 8023 8122: contig of 1594 bp in length
 8123 9244: contig of 1122 bp in length
 9245 10628: gap of unknown length
 10629 10728: contig of 1284 bp in length
 10729 11728: gap of unknown length
 11730 11829: contig of 1001 bp in length
 11830 13180: contig of 1351 bp in length
 13181 13280: gap of unknown length
 13281 14321: contig of 1041 bp in length
 14322 14421: gap of unknown length
 14422 15850: contig of 1429 bp in length
 15851 15950: gap of unknown length
 15951 17065: contig of 1115 bp in length
 17066 17165: gap of unknown length
 17166 18663: contig of 1498 bp in length
 18664 20471: contig of 1708 bp in length
 20472 20571: gap of unknown length
 20572 22115: contig of 1544 bp in length
 22116 22155: gap of unknown length
 22156 23583: contig of 1368 bp in length
 23584 23683: gap of unknown length
 23684 25100: contig of 1417 bp in length
 25101 25200: gap of unknown length
 25201 26828: contig of 1628 bp in length
 26829 26928: gap of unknown length
 26929 28052: contig of 1124 bp in length
 28053 28152: gap of unknown length
 28153 29437: contig of 1285 bp in length

29438 29537: gap of unknown length
29538 30900: contig of 1363 bp in length
30901 31000: gap of unknown length
31001 32025: contig of 1025 bp in length
32026 32125: gap of unknown length
32126 33494: contig of 1369 bp in length
33495 33594: gap of unknown length
33595 34746: contig of 1152 bp in length
34747 34846: gap of unknown length
34847 36376: contig of 1530 bp in length
36377 36476: gap of unknown length
36477 38351: contig of 1875 bp in length
38352 38451: gap of unknown length
38452 40679: contig of 2228 bp in length
40680 40779: gap of unknown length
40780 42712: contig of 1933 bp in length
42713 42812: gap of unknown length
42813 44002: contig of 1190 bp in length
44003 44102: gap of unknown length
44103 46067: contig of 1965 bp in length
46068 46167: gap of unknown length
46168 47905: contig of 1738 bp in length
47906 48005: gap of unknown length
48006 49836: contig of 1831 bp in length
49837 49936: gap of unknown length
49937 51790: contig of 1854 bp in length
51791 51890: gap of unknown length
51891 54730: contig of 2840 bp in length
54731 54830: gap of unknown length
54831 56607: contig of 1777 bp in length
56608 56707: gap of unknown length
56708 57790: contig of 1083 bp in length
57791 57890: gap of unknown length
57891 59527: contig of 1637 bp in length
59528 59627: gap of unknown length
59628 62056: contig of 2429 bp in length
62057 62156: gap of unknown length
62157 63464: contig of 1308 bp in length
63465 63564: gap of unknown length
63565 65003: contig of 1439 bp in length
65004 65103: gap of unknown length
65104 67035: contig of 1932 bp in length
67036 67135: gap of unknown length
67136 69383: contig of 2248 bp in length
69384 69483: gap of unknown length
69484 71920: contig of 2437 bp in length
71921 72020: gap of unknown length
72021 74582: contig of 2562 bp in length
74583 74682: gap of unknown length
74683 77272: contig of 2590 bp in length
77273 77372: gap of unknown length
77373 79400: contig of 2028 bp in length
79401 79500: gap of unknown length
79501 81759: contig of 2259 bp in length
81760 83356: contig of 1497 bp in length
83357 83456: gap of unknown length
83457 87053: contig of 3597 bp in length
87054 87153: gap of unknown length
87154 90859: contig of 3706 bp in length
90860 90959: gap of unknown length
90960 94572: contig of 3613 bp in length
94573 94672: gap of unknown length
94673 98285: contig of 3613 bp in length
98286 98385: gap of unknown length
98386 101083: contig of 2698 bp in length

Query Match 15.2% Score 20; DB 2; Length 187097;
Best Local Similarity 100.0%; Pred. No. 2;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 93 AGAGACCACAGAGCAGGAGA 112
|||||

Db 112893 AGAGACCACAGAGCAGGAGA 112874

RESULT 16

AC094105

LOCUS

DEFINITION

AC094105 192202 bp DNA linear HTG 15-NOV-2001;
Homo sapiens chromosome 5 clone RP11-468D11, WORKING DRAFT
SEQUENCE, 18 unordered pieces.

ACCESSION

AC094105

VERSION

AC094105.2 GI:16930895

KEYWORDS

HTG: HTGS_PHASE1; HTGS_DRAFT; HTGS_ACTIVEFIN.

SOURCE

Homo sapiens

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1 (bases 1 to 192202)
DOE Joint Genome Institute.

AUTHORS

Sequencing of Human Chromosome 5

JOURNAL

Unpublished

REFERENCE

2 (bases 1 to 192202)
DOE Joint Genome Institute.

AUTHORS

Direct Submission

JOURNAL

Submitted (14-SEP-2001) Production Sequencing Facility, DOE Joint
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
On Nov 15, 2001 this sequence version replaced gi:15617666.

COMMENT

-----Genome Center
Center: Joint Genome Institute
Center Code: JGI
Web site: http://www.jgi.doe.gov

Project Information
Center Project Name: 586160
Center clone name: RPCI-11_468D11

Summary Statistics
Consensus quality: 174989 bases at least Q40
Consensus quality: 184243 bases at least Q30
Consensus quality: 186010 bases at least Q20
Estimated insert size: 186280; agarose-fp estimation
Estimated insert size: 190502; sum-of-contigs estimation
Quality coverage: 6.08 in Q20 bases; agarose-fp estimation
Quality coverage: 5.95 in Q20 bases; sum-of-contigs estimation.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 18 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1
1154: contig of 1154 bp in length
1155 1254: gap of unknown length
1255 2278: contig of 1024 bp in length
2279 2378: gap of unknown length
2379 3567: contig of 1189 bp in length
3568 3667: gap of unknown length
3668 5532: contig of 1865 bp in length
5533 5632: gap of unknown length
5633 7425: contig of 1793 bp in length
7426 7525: gap of unknown length
7526 9566: contig of 2041 bp in length
9567 9666: gap of unknown length
9667 12444: contig of 2778 bp in length
12445 12544: gap of unknown length
12545 15344: contig of 2800 bp in length
15345 15444: gap of unknown length
15445 19247: contig of 3803 bp in length
19248 19247: gap of unknown length
19248 22855: contig of 3508 bp in length
22856 22956: gap of unknown length
22956 29417: contig of 6461 bp in length
29417 37476: contig of 7960 bp in length
37476 29517

* 37477 37576: gap of unknown length
 * 37577 49250: contig of 11674 bp in length
 * 49251 49350: gap of unknown length
 * 49351 65052: contig of 15702 bp in length
 * 65053 65152: gap of unknown length
 * 65153 81489: contig of 16337 bp in length
 * 81490 81589: gap of unknown length
 * 81590 100968: contig of 19379 bp in length
 * 100969 101068: gap of unknown length
 * 101069 144968: contig of 43900 bp in length
 * 144969 145068: gap of unknown length
 * 145069 192202: contig of 47134 bp in length.

FEATURES Location/Qualifiers
 source 1..192202

/organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="5"
 /clone="RP11-468D11"
 /clone_lib="RPCI human BAC library 11"
 55136 a 41193 c 40457 g 53685 t 1731 others

BASE COUNT

ORIGIN
 Query Match 15.2%; Score 20; DB 2; Length 192202;
 Best Local Similarity 100.0%; Pred. No. 2;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 64 GAGATGCAGAGAACTCT 83

Db 107984 GAGATGCAGAGAACTCT 108003

RESULT 17
 AC093570/c 204080 bp DNA linear HTG 14-AUG-2002
 LOCUS Mus musculus chromosome 5 clone rp23-115a12 strain C57BL/6J,
 DEFINITION WORKING DRAFT SEQUENCE, 3 ordered pieces.

AC093570

AC093570.21 GI:22218505

VERSION HTG; HTGS_PHASE2; HTGS_DRAFT.

KEYWORDS house mouse.

SOURCE Mus musculus

ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

AUTHORS Hu,X., Ornitz,D. and Roe,B.A.

TITLE 1 (bases 1 to 204080)

JOURNAL Mus musculus BAC Clone rp23-115a12

REFERENCE Unpublished

AUTHORS 2 (bases 1 to 204080)

TITLE Hu,X., Ornitz,D. and Roe,B.A.

JOURNAL Direct Submission

Submitted (05-SEP-2001) Department of Chemistry And Biochemistry,

The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,

OK 73019, USA

REFERENCE 3 (bases 1 to 204080)

AUTHORS Hu,X., Ornitz,D. and Roe,B.A.

TITLE Direct Submission

Submitted (14-AUG-2002) Department of Chemistry And Biochemistry,

The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,

OK 73019, USA

COMMENT On Aug 14, 2002 this sequence version replaced gi:21955040.

Center: Department Of Oklahoma

Center code:UOKNOR

----- Genome Center

----- Summary Statistics

Assembly program: XGAP4; version 4.5

Chemistry: Dye-terminator; 100% of reads

Consensus quality: 203878 bases at least Q40

Consensus quality: 204610 bases at least Q30

Consensus quality: 205218 bases at least Q20

Insert size: 205828; sum-of-contigs

Insert size: 215378; 4.9% error; agarose-fp

Quality coverage: 7.98x in Q20 bases; sum-of-contigs Quality

coverage: 7.77x in Q20 bases; agarose-fp

NOTE: This is a 'working draft' sequence. It currently

consists of 3 contigs. Gaps between the contigs

are represented as runs of N. The order of the pieces

is believed to be correct as given, however the sizes

of the gaps between them are based on estimates that have

provided by the submitter.

* the accession number will be preserved.
 * 1 12448: contig of 12448 bp in length
 * 12449 12548: gap of unknown length
 * 12549 65898: contig of 53350 bp in length
 * 65899 65998: gap of unknown length
 * 65999 204080: contig of 138082 bp in length.

FEATURES Location/Qualifiers
 source 1..204080

/organism="Mus musculus"

/strain="C57BL/6J"

/db_xref="taxon:10090"

/chromosome="5"

/clone="rp23-115a12"

/clone_lib="RPCI - 23 Female (C57BL/6J) Mouse BAC Library"

57304 a 45888 c 44284 g 56404 t 200 others

BASE COUNT

ORIGIN

Query Match 15.2%; Score 20; DB 2; Length 204080;
 Best Local Similarity 100.0%; Pred. No. 2;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 55 AAGCCACAGATGCAGAA 74

Db 133715 AAGCCACAGATGCAGAA 133696

RESULT 18

AL844896/c 207528 bp DNA linear HTG 29-JUL-2002

LOCUS Mus musculus chromosome 2 clone RP23-356L21, *** SEQUENCING IN

DEFINITION PROGRESS ***, 18 unordered pieces.

AL844896

AL844896.2 GI:22022653

VERSION HTG; HTGS_PHASE1.

KEYWORDS Mus musculus.

SOURCE Mus musculus

ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

AUTHORS 1 (bases 1 to 207528)

TITLE Direct Submission

Submitted (28-JUL-2002) Wellcome Trust Sanger Institute, Hinxton,

Cambridgeshire, CB10 1SA, UK. E-mail enquiries:

humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk

On Jul 30, 2002 this sequence version replaced gi:22003377.

Center: Genome Center

Center: Wellcome Trust Sanger Institute

Center code: SC

Web site: http://www.sanger.ac.uk

Contact: humquery@sanger.ac.uk

----- Project Information

Center project name: BM356L21

----- Summary Statistics

Assembly program: XGAP4; version 4.5

Chemistry: Dye-terminator; 100% of reads

Consensus quality: 203878 bases at least Q40

Consensus quality: 204610 bases at least Q30

Consensus quality: 205218 bases at least Q20

Insert size: 205828; sum-of-contigs

Insert size: 215378; 4.9% error; agarose-fp

Quality coverage: 7.98x in Q20 bases; sum-of-contigs Quality

coverage: 7.77x in Q20 bases; agarose-fp

NOTE: This is a 'working draft' sequence. It currently

consists of 18 contigs. The true order of the pieces

is not known and their order in this sequence record is

arbitrary. Gaps between the contigs are represented as

runs of N, but the exact sizes of the gaps are unknown.

This record will be updated with the finished sequence

as soon as it is available and the accession number will

be preserved.

1

8474 8573: gap of 100 bp

```

/note="assembly_fragment:01607
fragment_chain:2"
misc_feature
122517..140300
/note="assembly_fragment:01211
fragment_chain:2"
misc_feature
140401..148485
/note="assembly_fragment:00955
fragment_chain:3"
misc_feature
148586..161944
/note="assembly_fragment:01347
fragment_chain:3"
misc_feature
162045..171791
/note="assembly_fragment:01946
fragment_chain:3"
misc_feature
171892..180695
/note="assembly_fragment:02658
fragment_chain:3"
misc_feature
180796..202956
/note="assembly_fragment:02256
fragment_chain:3"
misc_feature
203057..207528
/note="assembly_fragment:02737
fragment_chain:3"

BASE COUNT 56479 a 47488 c 46773 g 55080 t 1708 others
ORIGIN

Query Match 15.2% Score 20: DB 2; Length 207528;
Best Local Similarity 100.0%; Pred. No. 2;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 105 GCAGGAGAGTGAGTGAGAA 124
|||||
Db 143847 GCAGGAGAGTGAGTGAGAA 143828

RESULT 19
AL844536/c
LOCUS
DEFINITION
AL844536
ACCESSION
AL844536.3 GI:22204804
VERSION
HTG; HTGS_PHASE1.
KEYWORDS
house mouse.
SOURCE
Mus musculus
ORGANISM
Eukaryota; Metazoa; Chordata; Craniala; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 241585)
Direct Submission
Submitted (02-AUG-2002) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
On Aug 11, 2002 this sequence version replaced gi:21955751.

COMMENT
----- Genome Center
Center: Wellcome trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: humquery@sanger.ac.uk
----- Project Information
Center project name: BM22A15
----- Summary Statistics
Assembly program: XGAP4; version 4.5
Chemistry: Dye-terminator; 100% of reads
Consensus quality: 236896 bases at least Q40
Consensus quality: 238326 bases at least Q30
Consensus quality: 239491 bases at least Q20
Insert size: 240585; sum-of-contigs
Insert size: 207539; 1.7% error; agarose-fp
Quality coverage: 4.96x in Q20 bases; sum-of-contigs Quality
coverage: 5.80x in Q20 bases; agarose-fp
-----
* NOTE: This is a 'working draft' sequence. It currently

```

* consists of 11 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

1 26384: contig of 26384 bp in length
 * 26385 26484: gap of 100 bp
 * 26485 38315: contig of 11831 bp in length
 * 38316 38415: gap of 100 bp
 * 38416 41944: contig of 3529 bp in length
 * 41945 42044: gap of 100 bp
 * 42045 56520: contig of 14476 bp in length
 * 56521 56620: gap of 100 bp
 * 56621 77754: contig of 21134 bp in length
 * 77755 77854: gap of 100 bp
 * 77855 96164: contig of 18310 bp in length
 * 96165 96264: gap of 100 bp
 * 96265 101917: contig of 5653 bp in length
 * 101918 102017: gap of 100 bp
 * 102018 146947: contig of 44930 bp in length
 * 146948 147047: gap of 100 bp
 * 147048 187542: contig of 40495 bp in length
 * 187543 187642: gap of 100 bp
 * 187643 192747: contig of 5105 bp in length
 * 192748 192847: gap of 100 bp
 * 192848 241585: contig of 48738 bp in length.

FEATURES

source
 1. 241585
 /organism="Mus musculus"
 /db_xref="taxon:10090"
 /chromosome="2"
 /clone="RP23-22A15"
 /clone_lib="RPCI-23"
 1. 26384
 /note="assembly_fragment:01220
 fragment_chain:1"
 26485..38315
 /note="assembly_fragment:01223
 fragment_chain:1"
 38416..41944
 /note="assembly_fragment:01396
 fragment_chain:1"
 42045..56520
 /note="assembly_fragment:01233
 fragment_chain:1"
 56621..77754
 /note="assembly_fragment:00572
 fragment_chain:2"
 77855..96164
 /note="assembly_fragment:01046
 fragment_chain:2"
 96265..101917
 /note="assembly_fragment:01569
 fragment_chain:2"
 102018..146947
 /note="assembly_fragment:00752
 fragment_chain:2"
 clone_end:T7
 vector_side:right
 147048..187542
 /note="assembly_fragment:00109
 fragment_chain:2"
 187643..192747
 /note="assembly_fragment:00452
 fragment_chain:2"
 192848..241585
 /note="assembly_fragment:00041
 fragment_chain:2"
 68487 a 53697 c 52331 g 66069 t 1001 others
 ORIGIN

Query Match 15.2%; Score 20; DB 2; Length 241585;
 Best Local Similarity 100.0%; Pred. No. 2;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 105 GCAGGAGAAAGTGAGTGA 124
 Db 59880 GCAGGAGAAAGTGAGTGA 59861
 RESULT 20
 AP000981/c 263050 bp DNA linear BCT 11-OCT-2001
 LOCUS Sulfolobus tokodaii genomic DNA, complete sequence, section:1/10.
 DEFINITION AP000981 BA000023
 ACCESSION AP000981.1 GI:15620945
 VERSION
 KEYWORDS
 SOURCE Sulfolobus tokodaii (strain:7) DNA.
 ORGANISM Sulfolobus tokodaii
 Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;
 Sulfolobus.
 REFERENCE
 1 Kavarabayasi,Y., Hino,Y., Horikawa,H., Jin-no,K., Takahashi,M.,
 Sekine,M., Baba,S., Ankai,A., Kosugi,H., Hosoyama,A., Fukui,S.,
 Nagai,Y., Nishijima,K., Otsuka,R., Nakazawa,H., Takamiya,M.,
 Kato,Y., Yoshizawa,T., Tanaka,T., Kudoh,Y., Yamazaki,J.,
 Kushida,N., Oguchi,A., Aoki,K., Masuda,S., Yanagii,M.,
 Nishimura,M., Yamagishi,A., Oshima,T. and Kikuchi,H.
 Complete genome sequence of an aerobic thermophilic
 crenarchaeon, Sulfolobus tokodaii strain7
 DNA Res. 8 (4), 123-140 (2001)
 2456156
 MEDLINE
 2 (bases 1 to 263050)
 Kavarabayasi,Y., Tanaka,T., Hino,Y. and Kikuchi,H.
 Direct Submission
 Submitted (22-DEC-1999) Yutaka Kavarabayasi, National Institute of
 Technology and Evaluation, Biotechnology Center, Nishihara 2-49-10,
 Shibuya-ku, Tokyo 151-0066, Japan (E-mail:kyutaka@nitech.go.jp,
 URL:htp://www.bio.nitech.go.jp/, Tel:81-3-3481-8951,
 Fax:81-3-3481-8424)
 COMMENT
 Kavarabayasi, Y. is officially affiliated with the National
 Institute of Advanced Industrial Science and Technology, Tsukuba,
 Ibaraki 305-0046, Japan
 Yamagishi, A. and Oshima, T. are at Tokyo University of Pharmacy
 and Life Science, Hachioji, Tokyo 192-0392, Japan
 The other authors are at the National Institute of Technology and
 Evaluation, Nishihara 2-49-10, Shibuya-ku, Tokyo 151-0066, Japan
 Homology analysis is performed by Smith-Waterman algorithm against
 Genbank release 109; EMBL release 56.0; SwissProt release 38.0;
 PIR-Protein release 62.0; and OWL release 31.4.
 E-mail address for comments and questions: kyutaka@nitech.go.jp
 ORF organization, sequence alignment and more information are
 available at W.W.W. site of Biotechnology Center,
 URL: http://www.bio.nitech.go.jp/dogan/genome_list-e.html/
 Location/Qualifiers
 1. 263050
 /organism="Sulfolobus tokodaii"
 /strain="7"
 /db_xref="taxon:111955"
 complement(526..1632)
 /gene="ST0001"
 complement(526..1632)
 /gene="ST0001"
 /codon_start=1
 /transl_table=11
 /product="368aa long hypothetical protein"
 /protein_id="BAB64942.1"
 /db_xref="GI:15620946"
 /translation="MSVIVYTVIDGELNVKQDQFFSNLATAVNLIQSFDLEPLK
 LHEIKYDGEADKDEVLPVMSKTFDILSVKFDLSDKFEFENYLGIVKYLIEKKIK
 ELIKKRGVEADKDEVLPVMSKTFDILSVKFDLSDKFEFENYLGIVKYLIEKKIK
 EELGYAPFLFTKPKRIGVDFYRIYVQLNEERFENYLGIVKYLIEKKIK
 LKSKTRLEIRNVSEFLNRYLLIYALRYHWILRTSKLTIGEAENVCLPILIASVDKPTQ
 FVKEIKDLYHRIENGVDSIRKEIENIGWYNITLPSRNIQSVREVNKNELFKIG

FEATURES

source
 1. 263050
 /organism="Sulfolobus tokodaii"
 /strain="7"
 /db_xref="taxon:111955"
 complement(526..1632)
 /gene="ST0001"
 complement(526..1632)
 /gene="ST0001"
 /codon_start=1
 /transl_table=11
 /product="368aa long hypothetical protein"
 /protein_id="BAB64942.1"
 /db_xref="GI:15620946"
 /translation="MSVIVYTVIDGELNVKQDQFFSNLATAVNLIQSFDLEPLK
 LHEIKYDGEADKDEVLPVMSKTFDILSVKFDLSDKFEFENYLGIVKYLIEKKIK
 ELIKKRGVEADKDEVLPVMSKTFDILSVKFDLSDKFEFENYLGIVKYLIEKKIK
 EELGYAPFLFTKPKRIGVDFYRIYVQLNEERFENYLGIVKYLIEKKIK
 LKSKTRLEIRNVSEFLNRYLLIYALRYHWILRTSKLTIGEAENVCLPILIASVDKPTQ
 FVKEIKDLYHRIENGVDSIRKEIENIGWYNITLPSRNIQSVREVNKNELFKIG

gene

MKLGPIWIVYLCMSIIVYKVGDFDKVPRV"
complement(1883. .2173)
/gene="STS001"
CDS
complement(1883. .2173)
/gene="STS001"
/note="PIR:C71147 percent identity: 36.264 in 91aa."
/codon_start=1
/transl_table=11
/product="96aa long conserved hypothetical protein"
/protein_id="BAB64943.1"
/db_xref="GI:15620947"
/translation="MEVSNVWKVVLTRLEQDAIKAVTIKEIYLPOLLTLNKSVN
LIERMEISVKEKIFIDYSLYIALERKSKLVTDKKQHEIAKKYISELI"
2263. .2496
/gene="STS002"
CDS
2263. .2496
/gene="STS002"
/note="OWL:A00972 percent identity: 28.070 in 57aa."
/codon_start=1
/transl_table=11
/product="77aa long conserved hypothetical protein"
/protein_id="BAB64944.1"
/db_xref="GI:15620948"
/translation="MRQLLSLSLTNSALLEVSGIWRISLIFFPKISSFFNSSIILL
IKFSSFLHLYFSSISRSNLSILKLYTFIL"

gene

complement(2268. .2537)
/gene="STS003"
CDS
complement(2268. .2537)
/gene="STS003"
/note="PIR:H69325 percent identity: 40.580 in 69aa."
/codon_start=1
/transl_table=11
/product="89aa long conserved hypothetical protein"
/protein_id="BAB64945.1"
/db_xref="GI:15620949"
/translation="MSKIETFFSVYVYVNMKVYSLRIDKELREMEKYNIKWNEI
ENFIRREELKEKELIKNEILQTMPTNSSAELVREDRNN"
complement(2722. .2937)
/gene="STS004"
CDS
complement(2722. .2937)
/gene="STS004"
/note="PIR:S74019 percent identity: 84.058 in 69aa."
/codon_start=1
/transl_table=11
/product="71aa long conserved hypothetical protein"
/protein_id="BAB64946.1"
/db_xref="GI:15620950"
/translation="MGKITGSDIDVILVIKRNDRNKALIDFFYEVERELGEKVSYL
FDVKVYDEKELPFKSLRDAVIK"
3144. .3401
/gene="STS005"
CDS
3144. .3401
/gene="STS005"
/note="Swiss Prot:P37413 percent identity: 52.000 in 25aa.
motif=prenyl group binding site (CAAX box)"
/codon_start=1
/transl_table=11
/product="85aa long hypothetical protein"
/protein_id="BAB64947.1"
/db_xref="GI:15620951"
/translation="MSFSLVILRKVAYPISKVYNKESNLIPWVLGGSPSLKSLGK
VVLKHLFLMPHKGLAEIIVLKLLSSPLPSSKILQCCMT"
complement(3247. .3399)
/gene="STS006"
CDS
complement(3247. .3399)
/gene="STS006"
/note="PIR:S74018 percent identity: 58.696 in 46aa."
/codon_start=1
/transl_table=11
/product="50aa long conserved hypothetical protein"
/protein_id="BAB64948.1"
/db_xref="GI:15620952"
/translation="WSYNIAEEFLRAKDYLKASELLFOOGLYEASALDSEVSARLYL

gene

NFLNLG"
3744. .5483
/gene="ST0002"
CDS
3744. .5483
/gene="ST0002"
/note="Swiss_Prot:P55577 percent identity: 37.004 in
227aa."
/codon_start=1
/transl_table=11
/product="579aa long hypothetical prolyl endopeptidase"
/protein_id="BAB64949.1"
/db_xref="GI:15620953"
/translation="MDEYEILENLSDPRTISFIEKENKETESKLGKKAMELYPLLLEM
DKEPYVLSMFAYDEENPAILLXGKSQLLGNKTIYIPPEGYVASEIWKVYNSKEIGV
SIERRGSDKIITLLISPEGKIRELGEMVSEPFYKGCELYIKSYRSPDPDGGDYPAD
RVFCDEIVYGDMKRPFEVTKVFDDELTKVQKGRYGEVYEGFDSLRKVDDEGE
VTDVIDFOQGEVYOKNNAYILGNTRKVEVDYVPLGVSHIGDKTAVRVEIKERTPLIF
YDIKKKIGEEVHDNTIFMDGKGTILFLETSEFNKRVVVKKSDGEGEVLMOQGVND
VITKDLIVKGDVLLHGFLVSKANNKPGIVYIGGFRIPLLPSTSVMRVLLNNGYSI
LITNLRGGYENGEEHMKAGMLNKNVFKDFAEFLRLVLMKGGKAVAMGSSNGLLVG
ATINEPYELIDCAVIGHPVLDMLRYDKLYVGYKVEEYGDNPDKYTEYLISYGFYHN
LKKGLPKTFVYTGINDDRVHPAHALYKVAKSKSLGNDVLMFLVNDSGHSIADPESKARE
ESYVVSFIECLR"
complement(5970. .8819)
/gene="ST0003"
CDS
complement(5970. .8819)
/gene="ST0003"
/note="Swiss_Prot:Q04069 percent identity: 32.836 in 67aa.
motif=ATP/GTP-binding site motif A (P-loop)"
/codon_start=1
/transl_table=11
/product="949aa long hypothetical protein"
/protein_id="BAB64950.1"
/db_xref="GI:15620954"
/translation="MIDCSSILSSLDLVQKFSQGVPTTTNNVKRDVLPITLCSRN
VRLDLNMQHOLEAINKLMOGKALVYNAETGGKTEIWNYSIALEMQKGFENVLAIV
PKALAGDQIERIKYIIEAGFSVSKQKTKIYEYIGVYIKYDGDVSNYIQSVAR
AKTLTNPEVVKNALQNHKISDLKKVRLIVVDEDFYFGSSKSTVLLHIIGIKDF
GIKQIVIMSATISDPVEVKPFDFEIIIGKSPANDTYIVLGKRDLELNAISKALNV
SPDDLKVFNFYASDPKKEIYQLQFDNAQLLPEYLDLKGCDLTIIFRSISEAN
NLVGLKGVNVDGTAPVAHVHSGIDKYVRQOVENDMRSGKLKVVVTVKTLQGIIDVG
NTRVYHVGIPDSVREFIOREGPKGMAKIKTESVIFPLSLSDAVLLEDYSTSKEW
LSLGSSELILLDPNEFLKIDVIRGILNDRFLAKSVGITGTLPOISYEQLHKTVPKL
LFDGRNCNVYDVRVVEKHQGLCDPMLNAIVRNSVINGRHHVETSDFSSSIM
CNNINISHVYNNATAYIEDICFSWKQHPDLNGDIERGKVMKVSVDLVLPFGDGGFKV
VHEIARKVFWYIESRRKIGQYKVEKIELEYTTPSNLKYDFLTYYVASELPDDVKNQ
DKGMFYLLALRLNYGIDLGLNFGVNSGILKWESEPTALLRRLMRKVKVVRGTE
LDCKLLDDVRNAQPPRRLELVLYDLPYTFNNFDFNEVKAIARFVHYLCNTIP1QIOL
QVISTGMPKSPILRSIILDSFGGKAVSSPTGEISLFDNEEAEVKAALKINIEFDA
IVPYGVSIGKKYFQNEVYIDKEINRLGGPITPSKFRLELNDLSLLKEENADN
SITRGEVEELTKIFRKRAETIRLMLNKAFLYTDK"
9382. .9549
/gene="STS007"
CDS
9382. .9549
/gene="STS007"
/note="OWL:AF072619 percent identity: 27.500 in 40aa."
/codon_start=1
/transl_table=11
/product="55aa long hypothetical cytochrome b"
/protein_id="BAB64951.1"
/db_xref="GI:15620955"
/translation="MKACTPSLLRENFIELIVMLTPISLRSSRRDLTLSPSRQEQSL
CILNIYELFYI"
complement(9819. .10127)
/gene="ST0004"

Query Match

15.2%; Score 20; DB 1; Length 263050;
Best Local Similarity 100.0%; Pred. No. 2;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 108 GGAGAGTGGAGTGAATTT 127
|||||
DB 68451 GGAGAGTGGAGTGAATTT 68432

RESULT 21

AP001760.

LOCUS

AP001760.

DEFINITION

Homo sapiens genomic DNA, chromosome 21q, section 104/105.

ACCESSION

AP001760.1 GI:7768687

VERSION

AP001760.1 GI:7768687

KEYWORDS

Homo sapiens DNA.

ORGANISM

Homo sapiens

REFERENCE

1 (sites)

AUTHORS

Hattori,M., Fujiyama,A., Taylor,T.D., Watanabe,H., Yada,T.,

Park,H.S., Toyoda,A., Ishii,K., Totoki,Y., Choi,D.K., Soeda,E.,

Ohki,M., Takagi,T., Sakaki,Y., Taudien,S., Blechschmidt,K.,

Polley,A., Menzel,U., Delabar,J., Kumpf,K., Lehmann,R.,

Patterson,D., Reichwald,K., Rump,A., Schillhabel,M., Schudy,A.,

Zimmermann,W., Rosenthal,A., Kudo,H., J., Shibuya,K., Kawasaki,K.,

Asakawa,S., Shintani,A., Sasaki,K., Nagamine,K., Mitsuyama,S.,

Antonarakis,S.E., Minoshima,S., Shimizu,N., Nordsiek,G.,

Hornischer,K., Barandt,P., Scharfe,M., Schoen,O., Desario,A.,

Reichert,J., Kauer,G., Bloeker,H., Ramser,J., Beck,A., Klages,S.,

Hennig,S., Riesselmann,L., Dagand,E., Wehrmeyer,S., Borzym,K.,

Gardiner,K., Nizetic,D., Francis,F., Leirach,H., Reinhardt,R. and

Yaspo,M.L.

TITLE

The DNA sequence of human chromosome 21

JOURNAL

Nature 405 (6784), 311-319 (2000)

MEDLINE

20289799

REFERENCE

2 (bases 1 to 340000)

AUTHORS

Hattori,M., Fujiyama,A., Taylor,T.D., Watanabe,H., Yada,T.,

Park,H.S., Toyoda,A., Ishii,K., Totoki,Y., Choi,D.K., Soeda,E.,

Ohki,M., Takagi,T., Sakaki,Y., Taudien,S., Blechschmidt,K.,

Polley,A., Menzel,U., Delabar,J., Kumpf,K., Lehmann,R.,

Patterson,D., Reichwald,K., Rump,A., Schillhabel,M., Schudy,A.,

Zimmermann,W., Rosenthal,A., Kudo,H., J., Shibuya,K., Kawasaki,K.,

Asakawa,S., Shintani,A., Sasaki,T., Nagamine,K., Mitsuyama,S.,

Antonarakis,S.E., Minoshima,S., Shimizu,N., Nordsiek,G.,

Hornischer,K., Barandt,P., Scharfe,M., Schoen,O., Desario,A.,

Reichert,J., Kauer,G., Bloeker,H., Ramser,J., Beck,A., Klages,S.,

Hennig,S., Riesselmann,L., Dagand,E., Wehrmeyer,S., Borzym,K.,

Gardiner,K., Nizetic,D., Francis,F., Leirach,H., Reinhardt,R. and

Yaspo,M.L.

TITLE

Direct Submission

JOURNAL

Submitted (10-APR-2000) The Chromosome 21 Mapping and Sequencing

COMMENT

Consortium: * RIKEN Genomic Sciences Center, Human Genome Research

Group * Institute of Molecular Biotechnology, Genome Analysis *

Keio University School of Medicine, Dept. of Molecular Biology *

GBF, Dept. of Genome Analysis * Max-Planck Institute for Molecular

Genetics (addresses see below)

On May 30, 2000 this sequence version replaced gi:7717454.

The chromosome 21 mapping and sequencing consortium consisting of

* RIKEN Genomic Sciences Center, Human Genome Research Group, *

Sagami-hara 228-8555, Japan,

* e-mail: hattori@sc.riken.go.jp

* URL: http://hgp.gsc.riken.go.jp/

and

* Institute of Molecular Biotechnology, Genome Analysis, *

Beutenbergstrasse 11, D-07745 Jena, Germany,

* e-mail: nshimizu@mb-med.keio.ac.jp

* URL: http://www.dmb.med.keio.ac.jp/

and

* Keio University School of Medicine, Molecular Biology, * Tokyo

160-8582, Japan,

* e-mail: nshimizu@mb-med.keio.ac.jp

* URL: http://www.dmb.med.keio.ac.jp/

and

* GBF, Dept. of Genome Analysis,

* Max-Planck Institute for Molecular Genetics,

* Ihnestrasse 73, D-14195 Berlin, Germany,

* e-mail: info-chr21@molgen.mpg.de

* URL: http://chr21.rz-berlin.mpg.de/

AL163305: Submitted (10-APR-2000).

Location/Qualifiers

1. 340000

/organism="Homo sapiens"

/db_xref="taxon:9606"

/chromosome="21"

<1. 21604

/organism="Homo sapiens"

/db_xref="taxon:9606"

/chromosome="21"

/map="21q22.3"

/clone="CTD2308H15, 5' partial"

/clone_lib="CTD BAC library"

/note="Accession No. AP000471"

21162. 59560

/organism="Homo sapiens"

/db_xref="taxon:9606"

/chromosome="21"

/map="21q22.3"

/clone="f50C10"

/clone_lib="CMF21 Fosmid library"

/note="Accession No. AP001477"

59460. 86378

/organism="Homo sapiens"

/db_xref="taxon:9606"

/chromosome="21"

/map="21q22.3"

/clone="f27E1"

/clone_lib="CMF21 Fosmid library"

/note="Accession No. AP000335"

85477. 121753

/organism="Homo sapiens"

/db_xref="taxon:9606"

/chromosome="21"

/map="21q22.3"

/clone="p3588"

/clone_lib="CMP21 P1 library"

/note="Accession No. AP000336"

121730. 175282

/organism="Homo sapiens"

/db_xref="taxon:9606"

/chromosome="21"

/map="21q22.3"

/clone="p1957"

/clone_lib="CMP21 P1 library"

/note="Accession No. AP000337"

174577. 256692

/organism="Homo sapiens"

/db_xref="taxon:9606"

/chromosome="21"

/map="21q22.3"

/clone="pP1D4"

/clone_lib="RP1 PAC library"

/note="Accession No. AP000338"

256594. 334367

/organism="Homo sapiens"

/db_xref="taxon:9606"

/chromosome="21"

/map="21q22.3"

/clone="PP9H11"

/clone_lib="RP1 PAC library"

/note="Accession No. AP000339"

333446. >340000

/organism="Homo sapiens"

/db_xref="taxon:9606"

/chromosome="21"

/map="21q22.3"

/clone="p1136, 3' partial"

/clone_lib="CMP21 P1 library"

```
repeat_region /note="Accession No. AP000340"
735..1036
/note="AluJb"
/rpt_family="SINE/Alu"
/rpt_type=DISPERSED
1437..1518
/note="AluSg/x"
/rpt_family="SINE/Alu"
/rpt_type=DISPERSED
complement(2134..2242)
/note="MER4B"
/rpt_family="LRR/MER4-group"
/rpt_type=DISPERSED
2409..2992
/note="Cpg island"
2905..124461
/genes="PCNT"
join(2905..3010,5105..5317,13125..13496,24856..24936,
25471..25726,26185..26240,27740..27914,28412..28548,
30157..30268,31832..32054,32715..32796,34181..34355,
35703..35920,42209..42663,45313..45370,45404..45868,
60423..60569,64561..64712,67471..67613,67928..68160,
69399..69561,69893..70105,75993..76222,76742..76858,
76963..76983,78318..78524,80279..80449,81059..81211,
89917..90795,91565..91720,94797..95567,96932..97034,
100698..100852,104559..104699,106350..106523,
107123..107318,108738..108960,109235..109385,
110257..110943,114631..114875,115706..115808,
116891..117064,118810..118929,119582..119811,
121224..121300,122537..122675,123421..123548,
124011..124461)
/genes="PCNT"
/note="pericentrin, kendrin (KIAA0402), Accession No.
U52962"
2905..3010
/genes="PCNT"
number=1
3261..3402
/note="(CGGGG)n"
/rpt_family="Simple_repeat"
/rpt_type=TANDEM
3391..3994
/genes="PCNT"
/note="PCNT"
/genes="Cpg island"
5105..5317
/genes="PCNT"
number=2
5460..5753
/note="AluSx"
/rpt_family="SINE/Alu"
/rpt_type=DISPERSED
complement(5779..6086)
/note="Alu"
/rpt_family="SINE/Alu"
/rpt_type=DISPERSED
complement(6127..6292)
/note="FRAM"
/rpt_family="SINE/Alu"
15.2%; Score 20; DB 9; Length 340000;
Query Match
Best Local Similarity 100.0%; Pred. No. 2;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 27 AATTGCCAGCTTGGATAAGG 46
|||||
Db 47589 AATTGCCAGCTTGGATAAGG 47608

RESULT 22
AE013648/c
LOCUS
DEFINITION
Yersinia pestis KIM section 48 of 415 of the complete genome.
ACCESSION AE013648 AE009952

AE013648.1 GI:21957165
Yersinia pestis KIM
Yersinia pestis KIM
Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
Yersinia.
1 (bases 1 to 11258)
Deng, W., Burland, V., Plunkett, G. III, Boutin, A., Mayhew, G. F.,
Liss, P., Perna, N. T., Rose, D. J., Mau, B., Zhou, S., Schwartz, D. C.,
Fetherston, J. D., Lindler, L. E., Brubaker, R. R., Plana, G. V.,
Straley, S. C., McDonough, K. A., Nilles, M. L., Matson, J. S.,
Blattner, F. R. and Perry, R. D.
Genome Sequence of Yersinia pestis KIM
J. Bacteriol. 184 (16), 4601-4611 (2002)
12142430
2 (bases 1 to 11258)
Deng, W., Burland, V., Plunkett, G. III, Boutin, A., Mayhew, G. F.,
Liss, P., Perna, N. T., Rose, D. J., Mau, B., Zhou, S., Schwartz, D. C.,
Fetherston, J. D., Lindler, L. E., Brubaker, R. R., Plana, G. V.,
Straley, S. C., McDonough, K. A., Nilles, M. L., Matson, J. S.,
Blattner, F. R. and Perry, R. D.
Direct Submission
Submitted (21-FEB-2002) Genetics, University of Wisconsin, 445
Henry Mall, Madison, WI 53706, USA
Location/Qualifiers
1..11258
/organism="Yersinia pestis KIM"
/strain="KIM"
/db_xref="taxon:187410"
85..204
/product="5S ribosomal RNA"
/function="rRNA; macromolecule synthesis; modification:
Ribosomal and stable RNAs"
666..1703
/genes="murB"
/note="y0471"
666..1703
/genes="murB"
/function="enzyme; murein sacculus, peptidoglycan"
/note="residues 7 to 345 of 345 are 61.94 pct identical to
residues 4 to 342 of 342 from E. coli K12 : B3972;
residues 7 to 345 of 345 are 62.53 pct identical to
residues 4 to 342 of 342 from GenPept : >gb|AA122970.1|
(AE008893) UDP-N-acetylenolpyruvylglucosamine reductase
[Salmonella typhimurium LT2]"
/codon_start=1
/transl_table=11
/product="UDP-N-acetylenolpyruvylglucosamine reductase"
/protein_id="AA084059.1"
/db_xref="GI:21957166"
/translation="MSNQRSLKHLNTFALPAYASNVISAGSVETLIAWHESKAKRQ
NMPGLENLALIPCGVSGAPIONIGAYVELKQVCEYVDLLDMDKGTQIRISAECQFG
YRDSIFKHYGNGFAIVSGIKLMSWPTLGYGDLHMDPLTVTATDIFNSVCTMR
SKLPDPMVTGNAGSPFNPVVSAAIAEEIVHCYPNAPHYLOPDGSVKLAAGLIDQCS
LRGYQIGGAHVHQOQALVLINQSEATGQDVHLARYIRQVQARFSIWLEPEVFIAD
NGEVNAVEHLS"
814..996
/genes="y0470"
814..996
/genes="y0470"
/function="unknown"
/note="residues 9 to 58 of 60 are 34.00 pct identical to
residues 247 to 295 of 436 from GenPept :
>gb|AAD22321.1|AC006955.7 (AC006955) hypothetical protein
[arabidopsis thaliana]"
/codon_start=1
/transl_table=11
/product="hypothetical"
/protein_id="AA084060.1"
/db_xref="GI:21957167"
/translation="MKVAMCYLLKISLGQYYSIGLWGLLPKTVLLGIYMSVLKIGI
NMVTHCKIICRGWKI"
```


REFERENCE
AUTHORS

- 1 (bases 1 to 12292)
da Silva, A.C.R., Ferro, J.A., Reinach, F.C., Farah, C.S., Furlan, L.R.,
Quaggio, R.B., Monteiro-Vitorello, C.B., Van Sluys, M.A., Almeida
Jr., N.F., Alves, L.M.C., do Amaral, A.M., Bertolini, M.C.,
Camargo, L.E.A., Camarotte, G., Cannavan, F., Cardozo, J.,
Chamberg, F., Ciapina, L.P., Cicarelli, R.M.B., Coutinho, L.L.,
Cursino-Santos, J.R., El-Dorry, H., Faria, J.B., Ferreira, A.J.S.,
Ferreira, R.C.C., Ferro, M.I.T., Formighieri, E.F., Franco, M.C.,
Greggio, C.C., Gruber, A., Katsuyama, A.M., Kishi, L.T., Leite
Jr., R.P., Lemos, E.G.M., Lemos, M.V.F., Locali, E.C., Machado, M.A.,
Madeira, A.M.B.N., Martinez-Rossi, N.M., Martins, E.C., Meidanis, J.,
Mench, C.F.M., Miyaki, C.Y., Moon, D.H., Moreira, L.M., Novo, M.T.M.,
Okura, V.K., Oliveira, M.C., Oliveira, V.R., Pereira Jr., H.A.,
Rosi, A., Sena, J.A.D., Silva, C., de Souza, R.F., Spinola, L.A.F.,
Takita, M.A., Tamura, R.E., Teixeira, E.C., Tezara, R.I.D., Trindade dos
Santos, M., Truffi, D., Tsai, S.M., White, F.F., Setubal, J.C. and
Kitajima, J.P.
- Comparison of the genomes of two Xanthomonas pathogens with
differing host specificities
Nature 417 (6887), 459-463 (2002)
22022145
12024217
- 2 (bases 1 to 12292)
da Silva, A.C.R., Ferro, J.A., Reinach, F.C., Farah, C.S., Furlan, L.R.,
Quaggio, R.B., Monteiro-Vitorello, C.B., Van Sluys, M.A., Almeida
Jr., N.F., Alves, L.M.C., do Amaral, A.M., Bertolini, M.C.,
Camargo, L.E.A., Camarotte, G., Cannavan, F., Cardozo, J.,
Chamberg, F., Ciapina, L.P., Cicarelli, R.M.B., Coutinho, L.L.,
Cursino-Santos, J.R., El-Dorry, H., Faria, J.B., Ferreira, A.J.S.,
Ferreira, R.C.C., Ferro, M.I.T., Formighieri, E.F., Franco, M.C.,
Greggio, C.C., Gruber, A., Katsuyama, A.M., Kishi, L.T., Leite
Jr., R.P., Lemos, E.G.M., Lemos, M.V.F., Locali, E.C., Machado, M.A.,
Madeira, A.M.B.N., Martinez-Rossi, N.M., Martins, E.C., Meidanis, J.,
Mench, C.F.M., Miyaki, C.Y., Moon, D.H., Moreira, L.M., Novo, M.T.M.,
Okura, V.K., Oliveira, M.C., Oliveira, V.R., Pereira Jr., H.A.,
Rosi, A., Sena, J.A.D., Silva, C., de Souza, R.F., Spinola, L.A.F.,
Takita, M.A., Tamura, R.E., Teixeira, E.C., Tezara, R.I.D., Trindade dos
Santos, M., Truffi, D., Tsai, S.M., White, F.F., Setubal, J.C. and
Kitajima, J.P.

TITLE

JOURNAL
MEDLINE

Nature 417 (6887), 459-463 (2002)

PUBMED

REFERENCE
AUTHORS

- 1 (bases 1 to 12292)
da Silva, A.C.R., Ferro, J.A., Reinach, F.C., Farah, C.S., Furlan, L.R.,
Quaggio, R.B., Monteiro-Vitorello, C.B., Van Sluys, M.A., Almeida
Jr., N.F., Alves, L.M.C., do Amaral, A.M., Bertolini, M.C.,
Camargo, L.E.A., Camarotte, G., Cannavan, F., Cardozo, J.,
Chamberg, F., Ciapina, L.P., Cicarelli, R.M.B., Coutinho, L.L.,
Cursino-Santos, J.R., El-Dorry, H., Faria, J.B., Ferreira, A.J.S.,
Ferreira, R.C.C., Ferro, M.I.T., Formighieri, E.F., Franco, M.C.,
Greggio, C.C., Gruber, A., Katsuyama, A.M., Kishi, L.T., Leite
Jr., R.P., Lemos, E.G.M., Lemos, M.V.F., Locali, E.C., Machado, M.A.,
Madeira, A.M.B.N., Martinez-Rossi, N.M., Martins, E.C., Meidanis, J.,
Mench, C.F.M., Miyaki, C.Y., Moon, D.H., Moreira, L.M., Novo, M.T.M.,
Okura, V.K., Oliveira, M.C., Oliveira, V.R., Pereira Jr., H.A.,
Rosi, A., Sena, J.A.D., Silva, C., de Souza, R.F., Spinola, L.A.F.,
Takita, M.A., Tamura, R.E., Teixeira, E.C., Tezara, R.I.D., Trindade dos
Santos, M., Truffi, D., Tsai, S.M., White, F.F., Setubal, J.C. and
Kitajima, J.P.
- Direct Submission
Submitted (28-NOV-2001) Departamento de Bioquímica, Universidade de
Sao Paulo, Av. Prof. Lineu Prestes 748, Sao Paulo, SP 05508-900,
Brazil

TITLE
JOURNALFEATURES
source

Location/Qualifiers
1. .12292
/organism="Xanthomonas axonopodis pv. citri str. 306"
/strain="306"
/db_xref="taxon:190486"
/note="pathovar: citri"
complement(215. .778)
/gene="adk"
/note="XAC3437"
complement(215. .778)
/gene="adk"
/note="Identified by sequence similarity; putative; ORF
located using Blastx/Glimmer/Genemark"
/codon_start=1
/transl_table=11
/product="adenylate kinase"
/protein_id="AAM38280.1"
/db_xref="GI:21109795"
/translation="MRLVLLGPPGSGKGTQATRLKDTDFPHISTGLLRAEVAAGSP
LGLKAEVARGDLVDDILLGLMLEALRGADYAKGFIIDGYPRNVAQNALDELICK
IQGLDARVQLDVASSELLVETAGAKAEGREDNPESVRKRLQVYDSTAPVIGFYE
QRGLKARVQGVSLDEVRISKALGR"
939. .2195
/gene="pfka"
/note="XAC3438"
939. .2195
/gene="pfka"
/note="Identified by sequence similarity; putative; ORF
located using Blastx/Glimmer/Genemark"
/codon_start=1

gene

CDS

gene
CDS
complement(527. .6006)
/gene="XAC3441"
/note="putative; ORF located using Glimmer/Genemark"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AAM38284.1"
/db_xref="GI:21109799"
/translation="MSRLARLLVLLLSLAQIAAATPPPTFFVDVADVPANGEGWEAF
IDLEQRDRDQNICDFTGCGDFSDYQPLRLRCVSHRVTGVSRVSCVITFGASEYSVD
PRSGYLRSDSRVWRC"APLKAAGTRLDMEYRTLAVANFLPEPLPGGAPPIDYGLICL"
6204. .6740
/gene="ppa"
/note="XAC3442"
6204. .6740
/gene="ppa"
/note="Identified by sequence similarity; putative; ORF
located using Blastx/Glimmer/Genemark"
/codon_start=1
/transl_table=11
/product="inorganic pyrophosphatase"
/protein_id="AAM38285.1"
/db_xref="GI:21109800"

gene

CDS

gene
CDS
complement(527. .6006)
/gene="XAC3441"
/note="putative; ORF located using Glimmer/Genemark"
/codon_start=1
/transl_table=11
/product="inorganic pyrophosphatase"
/protein_id="AAM38285.1"
/db_xref="GI:21109800"

/transl_table=11
/product="6-phosphofructokinase"
/protein_id="AAM38281.1"
/db_xref="GI:21109796"
/translation="WTGNLLYAQSGVTVATNATAGVTAETAEARAKIKVLAARNGIL
GALREELIDTSKESAAAIAAQAOTPGAGFSCRYKLSLEDSAKYERLLDLVLRADHV
RVFLYNGNSDADTAMKWSQAKAYGYPLCHIGVPTKIDNLAATDTCPGFSGSAKYT
VVSREALDVAAMDTSTKVFTEAMGRHAGLAAAGLAGADPGDPAQIILLPERA
YDQAFLAKVKOVVEKVCVVVASEGIDQAHGKFVADAGAAFGHAGLOJGGVSFL
AAQKQELGYKVVHTLPDYLORSARHLASKTDWEQAQVGAAYKAYAKGNNAVLPVI
ERYSDAPYRWKIVPAPLHKVANHEKMPSPFLKDKGFGITERARRYAPFLIKGEAALA
YSGDGLPKYVSLKNVAARAKLPAWEG"
complement(2444. .2920)
/gene="XAC3439"
complement(2444. .2920)
/gene="XAC3439"
/note="putative; ORF located using Glimmer/Genemark"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AAM38282.1"
/db_xref="GI:21109797"
/translation="MLTSQAFAADPTPELKPAPCTAQAQAVCAVHTLRQIPCAVRLG
VFTGNAQPYTLISVSRSSPTCQPRARFVDFAKAPSVASGVWFNDVIRVPSAACPQQ
AVTVWRKPADAKPQLDGGQSRILEDAKQQAAGKIPQVPMFAQMTMEGAQ"
3219. .5246
/gene="XAC3440"
3219. .5246
/gene="XAC3440"
/note="Identified by sequence similarity; putative; ORF
located using Blastx/Glimmer/Genemark"
/codon_start=1
/transl_table=11
/product="H+ translocating pyrophosphate synthase"
/protein_id="AAM38283.1"
/db_xref="GI:21109798"
/translation="MLEHYGLWALGCAVLAIYVIVSARVWVAOPSGNARMQETIAAA
IOEGARAYLNRYLTISVAGAVFLVGLFSLWYTAIFGALGVSLGLAGYIGNVSY
RANVTAEARHGIGKAMDVAFRGGAITGMVLVGLGVAGYAFVQLQGMGLPLEQNL
HALVGLFGSSLLSIFARLGGITFGADVGLGVKEAGIPDDPDRNPAVIADNVG
DNVDCAGNAADLFETVAVTVIATMLGSLTLADTGSNAVLYPLVIGSVSTIAITVGA
ARVYKDGSSINGALYKGVISGVLAALAYWPTQSLMRDNIHGATYALYACALIGLVL
TGLVWITEYTYGTQYTPVOHVASASTDGHGTNIAGLISMSKSTALPVIACVRAIWG
AFHFGGLYGIATAATAMLNAGMIVALDGYPTIDNAGGIAEMAELEPPEVNTIDPLD
AVGNTTKAVTKGYAIGSAALALVDFYHNLQAANPDQVYAFDLSDHTVILIGLIG
GLIPLYFGAMAEVAGAACAAVVEVRORFELPGINAGTAKPOYDRAVDMLTRSAIG
EMVPSLLPVVPIIYVLLGLLPRALGGLLIGTIVTGLFLAISMTGGGAWDNKKYIE
DGHFGKSGSEAKAATGDTGDPYKDTAGPAINPLIKINIIVALLVPLP"
complement(527. .6006)
/gene="XAC3441"
complement(527. .6006)
/gene="XAC3441"
/note="putative; ORF located using Glimmer/Genemark"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AAM38284.1"
/db_xref="GI:21109799"
/translation="MSRLARLLVLLLSLAQIAAATPPPTFFVDVADVPANGEGWEAF
IDLEQRDRDQNICDFTGCGDFSDYQPLRLRCVSHRVTGVSRVSCVITFGASEYSVD
PRSGYLRSDSRVWRC"APLKAAGTRLDMEYRTLAVANFLPEPLPGGAPPIDYGLICL"
6204. .6740
/gene="ppa"
/note="XAC3442"
6204. .6740
/gene="ppa"
/note="Identified by sequence similarity; putative; ORF
located using Blastx/Glimmer/Genemark"
/codon_start=1
/transl_table=11
/product="inorganic pyrophosphatase"
/protein_id="AAM38285.1"
/db_xref="GI:21109800"

gene /translation="MGLELVTSGKNLPPEINNVVIEIPKDSPEVKYEVKASGAIFVDR
ILSTPMRYPCNYGYVNTICGGDQADVLVLPPLPGSVVRCRPGVGLRMSDEAGS
DEKILAVIEKIFGYAHIEDINQVSSHMERIGHFHYKDKLKGKWKVLDGWCAGAA
EAKRLITVESVRYNSDAP"
6856 .8004
/gene="XAC3443"
6856 .8004
/gene="XAC3443"
/note="identified by sequence similarity; putative; ORF
located using Blastx/Glimmer/Genemark"
/codon_start=1
/transl_table=11
/product="response regulator"
/protein_id="AAM38286.1"
/db_xref="GI:21109801"
/translation="MSVTGDFWGSVLRILFVGDEASLPDSLVYVADLDGQAOQOVA
DNGSAIEAALSPPDVAIVPLDPTAATLGIQRTLRPDTRIALIDAQGGQFPP
ARITGVAHREPLMLAPVLEAVTSLELERELLSNPRRAAIGRIKPLSPPHLILS
LMHLEEDGADADIAKLIAGDPAIAAKVLIQLCNSAFFSGGRSTDLRTAVTRUGVA
TLRLVLASEVFSVOTLTPAERNAMORRALISRLAAKVLPTSAELGSTAALLADIG
LLLPVGRDEREAPAEAGDERLGHTEAGAYLLGLGMPPIIEAVAFHRRPLRSSMRSE
WVTGAVHATVATLASGESVDEEYLTQVGIARLPWSREQADTLGLTEA"
complement(8422..11268)
/gene="btub"
/note="XAC3444"
complement(8422..11268)
/gene="btub"
/note="identified by sequence similarity; putative; ORF
located using Blastx/Glimmer/Genemark"
/codon_start=1
/transl_table=11
/product="ionB-dependent receptor"
/protein_id="AAM38287.1"
/db_xref="GI:21109802"

gene

CDS

/translation="MNLRTSAVRLGLLPAGIATATPAFAAAQESTTTLDRLEITGS
RIRADVETQPIITLDRQALEKGLTSVADVLONLTSGSPALSRADLSGDEYVG
YYDRLNLGAQTLVLYNGKRLGSTGLQDLQIOPMSAERIEVLKDGASIIYGSDA
IAGVNNITRRFGDGAASVQFGQGGDDTQYFTVSGQGERGVTLISAEYSKQD
PVFAKNWFRSDGNSPNSPISQNGSWCNPLQVCDPDTAVQWTLNAGNPG
NNPADIHRITPAEYANQOMVQTIKERSVFANVNDVADITFNADYINERTTD
QIAGYPTQSGFCTPLGESSAFNPIGQDVFRRLEWEPRASENOLKTLRAPTLGS
YFDIGGTWMDVCLWNRNENTKRSRGMDSLIASRQALGPSFINGQVACQGTADP
ISLSACRPMPLLPYVAGQSLADPALQERLFTFTGLTKTTSFVANLSGSVEL
PAGDLGVAVGYEHRKEGRFSPDFAQSGESTGLGAKTQGIQSLDEYVLELNVLPLS

Query Match

Best Local Similarity 100.0%; Pred. No. 7.2; Length 12292;

Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 22 GAAGAAATGCCAGCTGG 40

Db 2168 GAAGAAATGCCAGCTGG 2186

RESULT 24

AC100719

LOCUS

DEFINITION

Mus musculus chromosome 1 clone RP23-170D17 map 1, LOW-PASS

SEQUENCE SAMPLING.

AC100719

AC100719.1 GI:17048085

HTG; HTGS PHASE0.

SOURCE

Mus musculus

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.

Birren,B., Linton,L., Nusbaum,C. and Lander,E.

Unpublished

TITLE

Mus musculus chromosome 1, clone RP23-170D17

REFERENCE

Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,

Anderson,S., Barna,N., Bastien,V., Boguslavskiy,L., Boukhvalter,B.,

Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B.,
Choepe,Y., Colangelo,M., Collins,S., Collamore,A., Cook,A.,
Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S.,
Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,
Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
Hagos,B., Headford,A., Horton,L., Hulme,W., Iliev,I., Johnson,R.,
Jones,C., Kamat,A., Karatas,A., Kells,C., LaRocque,K.,
Lamazares,R., Landers,T., Lehoczy,J., Levine,R., Liu,G.,
MacLean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C.,
McCarthy,M., McEwan,P., McKernan,K., MCPheeters,R., Meldrim,J.,
Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C.,
Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D.,
Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V.,
Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P.,
Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupbach,R.,
Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
Topham,K., Travers,M., Travis,N., Triglio,J., Vassiliev,H.,
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

Direct Submission

TITLE

JOURNAL

COMMENT

Submitted (22-NOV-2001) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
All repeats were identified using RepeatMasker:
http://ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu

----- Project Information
Center project name: L15998
Center clone name: 170_D_17

* NOTE: This record contains 83 individual
* sequencing reads that have not been assembled into
* contigs. Runs of N are used to separate the reads
* and the order in which they appear is completely
* arbitrary. Low-pass sequence sampling is useful for
* identifying clones that may be gene-rich and allows
* overlap relationships among clones to be deduced.
* However, it should not be assumed that this clone
* will be sequenced to completion. In the event that
* the record is updated, the accession number will
* be preserved.

1 733: contig of 733 bp in length
* 734 833: gap of 100 bp
* 834 1566: contig of 733 bp in length
* 1567 1666: gap of 100 bp
* 1667 2369: contig of 703 bp in length
* 2370 2469: gap of 100 bp
* 2470 3187: contig of 718 bp in length
* 3188 3287: gap of 100 bp
* 3288 3989: contig of 702 bp in length
* 3990 4089: gap of 100 bp
* 4090 4822: contig of 733 bp in length
* 4823 4922: gap of 100 bp
* 4923 5633: contig of 711 bp in length
* 5634 5733: gap of 100 bp
* 5734 6440: contig of 707 bp in length
* 6441 6540: gap of 100 bp
* 6541 7272: contig of 732 bp in length
* 7273 7372: gap of 100 bp
* 7373 8088: contig of 716 bp in length
* 8089 8188: gap of 100 bp
* 8189 8872: contig of 684 bp in length
* 8873 8972: gap of 100 bp
* 8973 9667: contig of 695 bp in length
* 9668 9767: gap of 100 bp
* 9768 10499: contig of 732 bp in length
* 10500 10599: gap of 100 bp
* 10600 11336: contig of 737 bp in length

* 11337 11436: gap of 100 bp
* 11437 12154: contig of 718 bp in length
* 12155 12254: gap of 100 bp
* 12255 13006: contig of 752 bp in length
* 13007 13106: gap of 100 bp
* 13107 13834: contig of 728 bp in length
* 13835 13934: gap of 100 bp
* 13935 14636: contig of 702 bp in length
* 14637 14736: gap of 100 bp
* 14737 15461: contig of 725 bp in length
* 15462 15561: gap of 100 bp
* 15562 16294: contig of 733 bp in length
* 16295 16394: gap of 100 bp
* 16395 17110: contig of 716 bp in length
* 17111 17210: gap of 100 bp
* 17211 17912: contig of 702 bp in length
* 17913 18012: gap of 100 bp
* 18013 18735: contig of 723 bp in length
* 18736 18835: gap of 100 bp
* 18836 19608: contig of 773 bp in length
* 19609 19708: gap of 100 bp
* 19709 20429: contig of 721 bp in length
* 20430 20529: gap of 100 bp
* 20530 21256: contig of 727 bp in length
* 21257 21356: gap of 100 bp
* 21357 22098: contig of 742 bp in length
* 22099 22198: gap of 100 bp
* 22199 22920: contig of 722 bp in length
* 22921 23020: gap of 100 bp
* 23021 23746: contig of 726 bp in length
* 23747 23846: gap of 100 bp
* 23847 24521: contig of 675 bp in length
* 24522 24621: gap of 100 bp
* 24622 25347: contig of 726 bp in length
* 25348 25447: gap of 100 bp
* 25448 26173: contig of 726 bp in length
* 26174 26273: gap of 100 bp
* 26274 27014: contig of 741 bp in length
* 27015 27114: gap of 100 bp
* 27115 27830: contig of 716 bp in length
* 27831 27930: gap of 100 bp
* 27931 28624: contig of 694 bp in length
* 28625 28724: gap of 100 bp
* 28725 29427: contig of 703 bp in length
* 29428 29527: gap of 100 bp
* 29528 30253: contig of 726 bp in length
* 30254 30353: gap of 100 bp
* 30354 31081: contig of 728 bp in length
* 31082 31181: gap of 100 bp
* 31182 31907: contig of 726 bp in length
* 31908 32007: gap of 100 bp
* 32008 32744: contig of 737 bp in length
* 32745 32844: gap of 100 bp
* 32845 33566: contig of 722 bp in length
* 33567 33666: gap of 100 bp
* 33667 34373: contig of 707 bp in length
* 34374 34473: gap of 100 bp
* 34474 35203: contig of 730 bp in length
* 35204 35303: gap of 100 bp
* 35304 36036: contig of 733 bp in length
* 36037 36136: gap of 100 bp
* 36137 36853: contig of 717 bp in length
* 36854 36953: gap of 100 bp
* 36954 37659: contig of 706 bp in length
* 37660 37759: gap of 100 bp
* 37760 38451: contig of 692 bp in length
* 38452 38551: gap of 100 bp
* 38552 39247: contig of 696 bp in length
* 39248 39347: gap of 100 bp
* 39348 40059: contig of 712 bp in length
* 40060 40159: gap of 100 bp
* 40160 40879: contig of 720 bp in length
* 40880 40979: gap of 100 bp

* 40980 41705: contig of 726 bp in length
* 41706 41805: gap of 100 bp
* 41806 42510: contig of 705 bp in length
* 42511 42610: gap of 100 bp
* 42611 43328: contig of 718 bp in length
* 43329 43428: gap of 100 bp
* 43429 44168: contig of 740 bp in length
* 44169 44268: gap of 100 bp
* 44269 44974: contig of 706 bp in length
* 44975 45074: gap of 100 bp
* 45075 45775: contig of 701 bp in length
* 45776 45875: gap of 100 bp
* 45876 46580: contig of 705 bp in length
* 46581 46680: gap of 100 bp
* 46681 47417: contig of 737 bp in length
* 47418 47517: gap of 100 bp
* 47518 48249: contig of 732 bp in length
* 48250 48349: gap of 100 bp
* 48350 49081: contig of 732 bp in length
* 49082 49181: gap of 100 bp
* 49182 49913: contig of 732 bp in length
* 49914 50013: gap of 100 bp
* 50014 50741: contig of 728 bp in length
* 50742 50841: gap of 100 bp
* 50842 51583: contig of 742 bp in length
* 51584 51683: gap of 100 bp
* 51684 52409: contig of 726 bp in length
* 52410 52509: gap of 100 bp
* 52510 53244: contig of 735 bp in length
* 53245 53344: gap of 100 bp
* 53345 54063: contig of 719 bp in length
* 54064 54163: gap of 100 bp
* 54164 54794: contig of 631 bp in length
* 54795 54894: gap of 100 bp
* 54895 55603: contig of 709 bp in length
* 55604 55703: gap of 100 bp
* 55704 56424: contig of 721 bp in length

Query Match 14.4%; Score 19; DB 2; Length 67972;
Best Local Similarity 100.0%; Pred. No. 7.5;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 43 AAGGCCAAGCTGAAGGCCA 61
|||||
Db 35409 AAGGCCAAGCTGAAGGCCA 35427

RESULT 25

AL606526/c

LOCUS

DEFINITION

Human DNA sequence from clone RP13-534F18 on chromosome 1, complete

ACCESSION

AL606526

VERSION

AL606526.10

KEYWORDS

HTG.

SOURCE

human.

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 113654)

Almeida, J.

Direct Submission

Submitted (15-NOV-2001) Wellcome Trust Sanger Institute, Hinxton,

Cambridgeshire, CB10 1SA, UK. E-mail enquiries:

humquerry@sanger.ac.uk

On Nov 16, 2001 this sequence version replaced gi:15808280.

During sequence assembly data is compared from overlapping clones.

Where differences are found these are annotated as variations

together with a note of the overlapping clone name. Note that the

variation annotation may not be found in the sequence submission

corresponding to the overlapping clone, as we submit sequences with

only a small overlap as described above.

This sequence was finished as follows unless otherwise noted: all

regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em: EMBL; SW: SWISSPROT; Tr: TrEMBL; Wp: WormPEP; Information on the WormPEP database can be found at <http://www.sanger.ac.uk/projects/C-elegans/wormpep> This sequence was generated from part of bacterial clone contigs of human chromosome 1, constructed by the Sanger Centre Chromosome 1 Mapping Group. Further information can be found at <http://www.sanger.ac.uk/HGP/Chr1>

RP13-534F18 is from the library RPCI-13.3 constructed by the group of Pieter de Jong. For further details see <http://www.chori.org/bacpac/home.htm>

VECTOR: pBAC3.6

IMPORTANT: This sequence is not the entire insert of clone RP13-534F18. It may be shorter because we sequence overlapping sections only once, except for a short overlap.

The true right end of clone RP13-534F18 is at 113654 in this sequence. The true right end of clone RP5-947P14 is at 2000 in this sequence.

FEATURES

Location/Qualifiers
1..113654
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="1"
/clone="RP13-534F18"
/clone_lib="RPCI-13.3"
39092 a 19726 c 20088 g 34748 t

BASE COUNT

ORIGIN
Query Match 14.4%; Score 19; DB 9; Length 113654;
Best Local Similarity 100.0%; Pred. NO. 7.6;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 112 AAGTGGAGTGAATTCCT 130
|||||
Db 46920 AAGTGGAGTGAATTCCT 46902

RESULT 26
AC079482
LOCUS
DEFINITION
AC079482.1 GI:9964847
VERSION
HTG; HTGS_PHASE1; HTGS_DRAFT.
KEYWORDS
Mus musculus.
SOURCE
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 121063)
Sequencing of Mouse
Unpublished
2 (bases 1 to 121063)
DOE Joint Genome Institute.
Direct Submission
Submitted (02-SEP-2000) Production Sequencing Facility, DOE Joint
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA

-----Genome Center
Center: Joint Genome Institute
Center Code: JGI
Web site: <http://www.jgi.doe.gov>

Project Information
Center Project Name: 1798378
Center clone name: RPCI-23_160C18

Summary Statistics
Consensus quality: 90928 bases at least Q40
Consensus quality: 105213 bases at least Q30
Consensus quality: 108646 bases at least Q20
Estimated insert size: 178000; agarose-fp estimation
Estimated insert size: 118063; sum-of-contigs estimation
Quality coverage: 8.18 in Q20 bases; agarose-fp estimation
Quality coverage: 12.33 in Q20 bases; sum-of-contigs estimation.
NOTE: This is a 'working draft' sequence. It currently
* consists of 31 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 1019: contig of 1019 bp in length
1020 1119: gap of unknown length
1120 2299: contig of 1180 bp in length
2300 2399: gap of unknown length
2400 3406: contig of 1007 bp in length
3407 3506: gap of unknown length
3507 4734: contig of 1228 bp in length
4735 4834: gap of unknown length
4835 6332: contig of 1498 bp in length
6333 6433: gap of unknown length
6434 7748: contig of 1316 bp in length
7749 9368: gap of unknown length
9369 9469: contig of 1520 bp in length
9470 11273: contig of 1805 bp in length
11274 11374: gap of unknown length
11375 12788: contig of 1415 bp in length
12789 12889: gap of unknown length
12890 14102: contig of 1114 bp in length
14103 15110: contig of 1608 bp in length
15111 15811: gap of unknown length
15812 17108: contig of 1298 bp in length
17109 17209: gap of unknown length
17210 18332: contig of 1124 bp in length
18333 18432: gap of unknown length
18433 20862: contig of 2430 bp in length
20863 20962: gap of unknown length
20963 23495: contig of 2533 bp in length
23496 23595: gap of unknown length
23596 25703: contig of 2108 bp in length
25704 25803: gap of unknown length
25804 27804: contig of 2001 bp in length
27805 27904: gap of unknown length
27905 32483: contig of 4579 bp in length
32484 32583: gap of unknown length
32584 33659: contig of 1086 bp in length
33660 33769: gap of unknown length
33770 36378: contig of 2609 bp in length
36379 36478: gap of unknown length
36479 43329: contig of 6851 bp in length
43330 43429: gap of unknown length
43430 49029: contig of 5600 bp in length
49030 49129: gap of unknown length
49130 55349: contig of 6220 bp in length
55350 61174: contig of 5725 bp in length
61175 61274: gap of unknown length
61275 66048: contig of 4774 bp in length
66049 66148: gap of unknown length
66149 74177: contig of 8029 bp in length
74178 74277: gap of unknown length
74278 80199: contig of 5922 bp in length
80200 80299: gap of unknown length
80300 87977: contig of 7678 bp in length
87978 88077: gap of unknown length

```
* 88078 95183: contig of 7106 bp in length
* 95184 95283: gap of unknown length
* 95284 106366: contig of 11083 bp in length
* 106367 106466: gap of unknown length
* 106467 121063: contig of 14597 bp in length.
FEATURES
    source
        Location/Qualifiers
            1..121063
                /organism="Mus musculus"
                /db_xref="taxon:10090"
                /clone="RP23-160C18"
                /clone_lib="RPCI mouse BAC library 23"
                /clone_23="25574 g 32817 t 3071 others"
BASE COUNT 33621 a 25980 c 25574 g 32817 t 3071 others
ORIGIN
    Query Match 14.4%; Score 19; DB 2; Length 121063;
    Best Local Similarity 100.0%; Pred. NO. 7.6;
    Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 43 AAGGCCAAGCTGAAGGCCA 61
Db 55036 AAGGCCAAGCTGAAGGCCA 55054

RESULT 27
AC079482/c
LOCUS AC079482 121063 bp DNA linear HTG 02-SEP-2000
DEFINITION Mus musculus clone RP23-160C18, WORKING DRAFT SEQUENCE, 31
unordered pieces.
ACCESSION AC079482
VERSION AC079482.1 GI:9964847
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE Mus musculus.
ORGANISM Mus musculus
    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
    Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
    1 (bases 1 to 121063)
    DOE Joint Genome Institute.
    Sequencing of Mouse
    Unpublished
    JOURNAL
    TITLE
    REFERENCE
    AUTHORS
    TITLE
    JOURNAL
    COMMENT
    Center: Joint Genome Institute
    Center Code: JGI
    Web site: http://www.jgi.doe.gov
    -----
    Project Information
    Center Project Name: 1798378
    Center clone name: RPCI-23_160C18
    -----
    Summary Statistics
    Consensus quality: 90928 bases at least Q40
    Consensus quality: 105213 bases at least Q30
    Consensus quality: 108646 bases at least Q20
    Estimated insert size: 178000; agarose-fp estimation
    Quality coverage: 8.18 in Q20 bases; sum-of-contigs estimation
    Quality coverage: 12.33 in Q20 bases; sum-of-contigs estimation.
    * NOTE: This is a 'working draft' sequence. It currently
    * consists of 31 contigs. The true order of the pieces
    * is not known and their order in this sequence record is
    * arbitrary. Gaps between the contigs are represented as
    * runs of N, but the exact sizes of the gaps are unknown.
    * This record will be updated with the finished sequence
    * as soon as it is available and the accession number will
    * be preserved.
    1 1019: contig of 1019 bp in length
    * 1020 1119: gap of unknown length
    * 1120 2299: contig of 1180 bp in length
    * 2300 2399: gap of unknown length

3406: contig of 1007 bp in length
3506: gap of unknown length
4734: contig of 1228 bp in length
4834: gap of unknown length
6332: contig of 1498 bp in length
6432: gap of unknown length
7748: contig of 1316 bp in length
7848: gap of unknown length
9368: gap of unknown length
9468: gap of unknown length
11273: contig of 1805 bp in length
11373: gap of unknown length
11374 12788: contig of 1415 bp in length
12789 12888: gap of unknown length
14002: contig of 1114 bp in length
14102: gap of unknown length
15102: contig of 1608 bp in length
15810: gap of unknown length
15811 17108: contig of 1298 bp in length
17109 17208: gap of unknown length
18332: contig of 1124 bp in length
18432: gap of unknown length
20862: contig of 2430 bp in length
20962: gap of unknown length
23495: contig of 2533 bp in length
23496 23595: gap of unknown length
25703: contig of 2108 bp in length
25803: gap of unknown length
27804: contig of 2001 bp in length
27904: gap of unknown length
32483: contig of 4579 bp in length
32484 32583: gap of unknown length
32584 33669: contig of 1086 bp in length
33670 33769: gap of unknown length
36378: contig of 2609 bp in length
36379 36478: gap of unknown length
43329: contig of 6851 bp in length
43429: gap of unknown length
49029: contig of 5600 bp in length
49129: gap of unknown length
53449: contig of 6220 bp in length
5350 55449: gap of unknown length
61174: contig of 5725 bp in length
61175 61274: gap of unknown length
66048: contig of 4774 bp in length
66049 66148: gap of unknown length
74177: contig of 8029 bp in length
74178 74277: gap of unknown length
80199: contig of 5922 bp in length
80200 80299: gap of unknown length
87977: contig of 7678 bp in length
88077: gap of unknown length
95183: contig of 7106 bp in length
95184 95283: gap of unknown length
95284 106366: contig of 11083 bp in length
106367 121063: contig of 14597 bp in length.
FEATURES
    source
        Location/Qualifiers
            1..121063
                /organism="Mus musculus"
                /db_xref="taxon:10090"
                /clone="RP23-160C18"
                /clone_lib="RPCI mouse BAC library 23"
                /clone_23="25574 g 32817 t 3071 others"
BASE COUNT 33621 a 25980 c 25574 g 32817 t 3071 others
ORIGIN
    Query Match 14.4%; Score 19; DB 2; Length 121063;
    Best Local Similarity 100.0%; Pred. NO. 7.6;
    Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 43 AAGGCCAAGCTGAAGGCCA 61
Db 66960 AAGGCCAAGCTGAAGGCCA 66942
```

COMMENT

On Apr 7, 2002 this sequence version replaced gi:17974075. During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em: EMBL; Sw: SWISSPROT; Tr: TREMBL; Wp: WORMPEP; Information on the WORMPEP database can be found at <http://www.sanger.ac.uk/Projects/C-elegans/wormpep> RP23-330D1 is from the RPCI-23 Mouse PAC Library constructed by the group of Pieter de Jong. For further details see <http://www.chori.org/bacpac/home.htm>

VECTOR: pBACE3.6

----- Genome Center
Center: UK Medical Research Council
Center code: UK-MRC
Web site: <http://mrcseq.har.mrc.ac.uk>
Contact: mrcseq@har.mrc.ac.uk

FEATURES

source
1..137782
/organism="Mus musculus"
/db_xref="taxon:10090"
/chromosome="13"
/clone="RP23-330D1"
/clone.lib="RPCI-23"

BASE COUNT 40627 a 27285 c 27611 g 42259 t
ORIGIN

Query Match 14.4%; Score 19; DB 10; Length 137782;
Best Local Similarity 100.0%; Pred. No. 7.6;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 43 AAGCCCAAGCTGAAGGCCA 61
|||||
Db 136870 AAGCCCAAGCTGAAGGCCA 136888

RESULT 30
AC125887 155726 bp DNA linear HTG 23-JUL-2002

LOCUS
DEFINITION
Rattus norvegicus clone CH230-274J3, *** SEQUENCING IN PROGRESS
***, 74 unordered pieces.

ACCESSION
AC125887
VERSION
AC125887.1 GI:21671483

KEYWORDS
HTG; HTGS_PHASE1.
SOURCE
Rattus norvegicus

ORGANISM
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.

REFERENCE
1 (bases 1 to 155726)
Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,
Alsbrooks,S.L., Amarantunge,H.C., Are,J.R., Ayale,M., Banks,T.,
Barbaria,J., Benton,J., Bimage,K., Blankenburg,K., Bonnin,D.,
Bouck,J., Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P.,
Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,
Carroll,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,
Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,
Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,
Devilla,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,
Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,
Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,

RESULT 28

AC008570/c

LOCUS

DEFINITION

AC008570

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

127392 bp DNA linear PRI 29-SEP-2001

127392 bp DNA linear PRI 29-SEP-2001

127392 bp DNA linear PRI 29-SEP-2001

127392 bp DNA linear PRI 29-SEP-2001

127392 bp DNA linear PRI 29-SEP-2001

127392 bp DNA linear PRI 29-SEP-2001

127392 bp DNA linear PRI 29-SEP-2001

127392 bp DNA linear PRI 29-SEP-2001

127392 bp DNA linear PRI 29-SEP-2001

127392 bp DNA linear PRI 29-SEP-2001

127392 bp DNA linear PRI 29-SEP-2001

127392 bp DNA linear PRI 29-SEP-2001

127392 bp DNA linear PRI 29-SEP-2001

127392 bp DNA linear PRI 29-SEP-2001

127392 bp DNA linear PRI 29-SEP-2001

127392 bp DNA linear PRI 29-SEP-2001

127392 bp DNA linear PRI 29-SEP-2001

127392 bp DNA linear PRI 29-SEP-2001

127392 bp DNA linear PRI 29-SEP-2001

127392 bp DNA linear PRI 29-SEP-2001

127392 bp DNA linear PRI 29-SEP-2001

127392 bp DNA linear PRI 29-SEP-2001

127392 bp DNA linear PRI 29-SEP-2001

127392 bp DNA linear PRI 29-SEP-2001

127392 bp DNA linear PRI 29-SEP-2001

127392 bp DNA linear PRI 29-SEP-2001

127392 bp DNA linear PRI 29-SEP-2001

127392 bp DNA linear PRI 29-SEP-2001

127392 bp DNA linear PRI 29-SEP-2001

127392 bp DNA linear PRI 29-SEP-2001

127392 bp DNA linear PRI 29-SEP-2001

127392 bp DNA linear PRI 29-SEP-2001

127392 bp DNA linear PRI 29-SEP-2001

127392 bp DNA linear PRI 29-SEP-2001

127392 bp DNA linear PRI 29-SEP-2001

127392 bp DNA linear PRI 29-SEP-2001

127392 bp DNA linear PRI 29-SEP-2001

127392 bp DNA linear PRI 29-SEP-2001

127392 bp DNA linear PRI 29-SEP-2001

127392 bp DNA linear PRI 29-SEP-2001

127392 bp DNA linear PRI 29-SEP-2001

127392 bp DNA linear PRI 29-SEP-2001

127392 bp DNA linear PRI 29-SEP-2001

127392 bp DNA linear PRI 29-SEP-2001

127392 bp DNA linear PRI 29-SEP-2001

127392 bp DNA linear PRI 29-SEP-2001

127392 bp DNA linear PRI 29-SEP-2001

127392 bp DNA linear PRI 29-SEP-2001

127392 bp DNA linear PRI 29-SEP-2001

127392 bp DNA linear PRI 29-SEP-2001

127392 bp DNA linear PRI 29-SEP-2001

127392 bp DNA linear PRI 29-SEP-2001

127392 bp DNA linear PRI 29-SEP-2001

Earnhart, C., Edgar, D., Edwards, C.C., Elhaj, C., Escotto, M., Falls, T., Ferraguto, D., Flagg, N., Ford, J., Foster, P., Frantz, P., Gabisi, A., Gao, J., Garcia, A., Garner, T., Garza, N., Gill, R., Gorrell, J.H., Guevara, W., Gunaratne, P., Hale, S., Hamilton, K., Harris, C., Harris, K., Hart, M., Havlak, P., Hawes, A., Hernandez, J., Hernandez, O., Hodgson, A., Hogue, M., Holloway, C., Hollins, B., Homsi, F., Howard, S., Huber, J., Hulyk, S., Hume, J., Jackson, L.E., Jacobson, B., Jia, Y., Johnson, R., Jolivet, S., Joudah, S., Karlsson, E., Kelly, S., Khan, U., King, L., Korvah, J., Kovar, C., Kratovic, J., Kureshi, A., Landry, N., Leal, B., Lewis, L.C., Lewis, L., Li, J., Li, Z., Lichtarge, O., Lieu, C., Liu, R., Liu, W., Loulseghe, H., Lozano, R.J., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J., Maheshwari, M., Mapa, P., Martin, R., Martindale, A., Martinez, E., Massey, E., Mawhney, E., McLeod, M.P., Meador, M., Mei, G., Metzger, M., Miner, G., Miner, Z., Mitchell, T., Mohabbat, K., Morgan, M., Morris, S., Moser, M., Neal, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N., Nguyen, N., Nickerson, E., Nwokenkwo, S., Ogih, M., Okwuonu, G., Oragunye, N., Oviedo, R., Pace, A., Payton, B., Peery, J., Perez, L., Peters, L., Pickens, R., Primus, E., Pu, L.L., Quiles, M., Ren, Y., Rives, M., Rojas, A., Rojibokan, I., Rolfe, M., Ruiz, S., Savary, G., Scherer, S., Scott, G., Shen, H., Shoostari, N., Sisson, I., Sodergren, E., Sonaike, T., Sparks, A., Stanley, H., Stone, H., Sutton, A., Vatek, A., Tabor, P., Tamerisa, A., Tamerisa, K., Tang, H., Tansey, J., Taylor, C., Taylor, T., Telford, B., Thomas, N., Thomas, S., Usmani, K., Vasquez, L., Vera, V., Villalon, D., Vinson, R., Wang, Q., Wang, S., Ward-Moore, S., Warren, R., Washington, C., Watlington, S., Williams, G., Williamson, A., Wleczky, R., Wooden, S., Worley, K., Wu, C., Wu, Y., Wu, Y.F., Zhou, J., Zorrilla, S., Nelson, D., Weinstock, G., and Gibbs, R.

Direct Submission
Unpublished
2 (bases 1 to 155726)
Worley, K.C.
Direct Submission
Submitted (02-JUL-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 155726)
Worley, K.C.
Direct Submission
Submitted (23-JUL-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu

Project Information
Center project name: GWLD
Center clone name: CH230-274J3

Summary Statistics
Sequencing vector: Plasmid
Chemistry: Dye-terminator Big Dye: 100% of reads
Assembly program: Phrap: version 0.990329
Consensus quality: 93488 bases at least Q40
Consensus quality: 99060 bases at least Q30
Consensus quality: 102929 bases at least Q20

* NOTE: Estimated insert size may differ from sequence length
(see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently consists of 74 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.
* 1 1065: contig of 1065 bp in length
* 1066 1165: gap of unknown length
* 1166 2186: contig of 1021 bp in length
* 2187 2286: gap of unknown length

2287 3495: contig of 1209 bp in length
3496 3595: gap of unknown length
3596 4549: contig of 1054 bp in length
4550 4749: gap of unknown length
4750 5775: contig of 1026 bp in length
5776 5875: gap of unknown length
5876 7013: contig of 1138 bp in length
7014 8585: contig of 1472 bp in length
8586 8686: gap of unknown length
8687 9693: contig of 1007 bp in length
9694 11295: contig of 1503 bp in length
11296 12652: contig of 1257 bp in length
12653 12752: gap of unknown length
12753 14036: contig of 1284 bp in length
14037 14136: gap of unknown length
14138 15653: contig of 1517 bp in length
15654 15753: gap of unknown length
15754 17043: contig of 1290 bp in length
17044 17143: gap of unknown length
17144 18447: contig of 1304 bp in length
18448 18547: gap of unknown length
18549 20191: contig of 1643 bp in length
20192 20290: gap of unknown length
20291 21436: contig of 1146 bp in length
21437 21536: gap of unknown length
21537 23361: contig of 1825 bp in length
23362 23461: gap of unknown length
23462 24927: contig of 1466 bp in length
24928 25027: gap of unknown length
25028 26574: contig of 1547 bp in length
26575 27768: contig of 1094 bp in length
27769 27868: gap of unknown length
27869 29089: contig of 1221 bp in length
29090 30291: gap of unknown length
30292 30391: gap of unknown length
30392 31632: contig of 1241 bp in length
31633 31732: gap of unknown length
31733 33206: contig of 1474 bp in length
33207 33306: gap of unknown length
33307 34489: contig of 1183 bp in length
34490 36149: gap of unknown length
36150 36249: gap of unknown length
36250 38072: contig of 1823 bp in length
38073 39507: contig of 1335 bp in length
39508 39607: gap of unknown length
39609 41207: contig of 1600 bp in length
41208 41307: gap of unknown length
41308 43733: contig of 2426 bp in length
43734 43833: gap of unknown length
43834 45306: contig of 1473 bp in length
45307 45406: gap of unknown length
45407 46935: contig of 1529 bp in length
46936 47035: gap of unknown length
47036 48280: contig of 1245 bp in length
48281 48381: gap of unknown length
48382 50905: contig of 2525 bp in length
50906 51005: gap of unknown length
51006 52013: contig of 1008 bp in length
52014 52113: gap of unknown length
52114 53357: contig of 1244 bp in length
53358 53457: gap of unknown length
53458 54671: contig of 1214 bp in length
54672 54771: gap of unknown length
54772 56149: contig of 1378 bp in length
56150 56249: gap of unknown length
56250 58450: contig of 2201 bp in length

TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

* 58451 58550: gap of unknown length
 * 58551 58702: contig of 1152 bp in length
 * 59703 59802: gap of unknown length
 * 59803 61210: contig of 1408 bp in length
 * 61211 61310: gap of unknown length
 * 61311 62734: contig of 1424 bp in length
 * 62735 62834: gap of unknown length
 * 62835 65673: contig of 2839 bp in length
 * 65674 65774: gap of unknown length
 * 65775 67274: contig of 1501 bp in length
 * 67275 67375: gap of unknown length
 * 67375 69075: contig of 1701 bp in length
 * 69076 69175: gap of unknown length
 * 69176 70561: contig of 1385 bp in length
 * 70561 70661: gap of unknown length
 * 70661 72979: contig of 2318 bp in length
 * 72979 73079: gap of unknown length
 * 73079 75785: contig of 2707 bp in length
 * 75786 75885: gap of unknown length
 * 75886 77423: contig of 1537 bp in length
 * 77423 77523: gap of unknown length
 * 77523 79027: contig of 1505 bp in length
 * 79028 79128: gap of unknown length
 * 79128 81056: contig of 1929 bp in length
 * 81057 81157: gap of unknown length
 * 81157 83211: contig of 2055 bp in length
 * 83212 83312: gap of unknown length
 * 83312 85650: contig of 2339 bp in length
 * 85651 85750: gap of unknown length
 * 85751 87045: contig of 1294 bp in length
 * 87045 87144: gap of unknown length

Query Match 14.4%; Score 19; DB 2; Length 155726;

Best Local Similarity 100.0%; Pred. No. 7.7;

Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 94 GAGACCACGAGCAGGAGA 112
 Db 102872 GAGACCACGAGCAGGAGA 102890
 |||||

RESULT 31

AL645923

LOCUS

DEFINITION Mouse DNA sequence from clone RP23-75H1 on chromosome 13, complete
 sequence.

ACCESSION

AL645923

VERSION

GI:19572045

KEYWORDS

HTG.

SOURCE

house mouse.

ORGANISM

Mus musculus

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Submitted (29-JUN-2002) Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
 On Mar 21, 2002 this sequence version replaced gi:19335835.
 During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.
 This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. The following

abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information on the WORMPEP database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep RP23-75H1 is from the RPCI-23 Mouse PAC Library
 constructed by the group of Pieter de Jong.
 For further details see <http://www.chori.org/bacpac/home.htm>
 VECTOR: pBACE3.6

----- Genome Center

Center: UK Medical Research Council

Center code: UK-MRC

Web site: <http://mrcseq.har.mrc.ac.uk>

Contact: mouse@har.mrc.ac.uk

FEATURES

source

Location/Qualifiers

1. 157361

/organism="Mus musculus"

/db_xref="taxon:10090"

/chromosome="13"

/clone="RP23-75H1"

/clone_lib="RPCI-23"

BASE COUNT 45215 a 31432 c 31628 g 49086 t

ORIGIN

Query Match 14.4%; Score 19; DB 10; Length 157361;

Best Local Similarity 100.0%; Pred. No. 7.7;

Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 43 AAGCCCAAGCTGAAGGCCA 61

Db 82294 AAGCCCAAGCTGAAGGCCA 82312

|||||

AC111338

LOCUS

DEFINITION

Rattus norvegicus clone CH230-240A17, *** SEQUENCING IN PROGRESS

***, 68 unordered pieces.

ACCESSION

AC111338

VERSION

GI:21735860

KEYWORDS

HTG; HTGS, PHASE1.

SOURCE

Norway rat.

ORGANISM

Rattus norvegicus

REFERENCE

1 (bases 1 to 158286)

AUTHORS

Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,

Alsbrooks,S.L., Amaratunge,H.C., Are,J.R., Ayele,M., Banks,T.,

Barbieri,J., Benton,J., Binage,K., Blankenburg,K., Bonnin,D.,

Bouck,J., Bowie,S., Brivea,M., Brown,E., Brown,M., Bryant,N.P.,

Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,

Carroll,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chaves,D.,

Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,

Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,

Devilay,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,

Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,

Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,

Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,

Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,

Gabriel,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,

Gorell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,

Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J.,

Hernandez,O., Hodgson,A., Hogues,M., Holloway,C., Hollins,B.,

Homsí,F., Howard,S., Huber,J., Hulyk,S., Hume,J., Jackson,L.E.,

Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S.,

Karlsson,E., Kelly,S., Khan,U., King,L., Korvan,J., Kovar,C.,

Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.,

Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W., Loulseged,H.,

Lozadó,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,

Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E.,

Massey,E., Mawhney,E., McLeod,M.P., Meador,M., Mei,G., Metzker,M.,

Miner, G., Miner, Z., Mitchell, T., Mohabbat, K., Morgan, M., Morris, S., Moser, M., Neal, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N., Nguyen, N., Nickerson, E., Nwokenkwo, S., Ogih, M., Okwuonu, G., Oragunye, N., Oviedo, R., Pace, A., Payton, B., Peery, J., Perez, L., Peters, L., Pickens, R., Primus, E., Pui, L., Quiles, M., Ren, Y., Rives, M., Rojas, A., Rojubokan, I., Rolfe, M., Ruiz, S., Savary, G., Scher, S., Scott, G., Shen, H., Shooshtari, N., Sisson, I., Sodergren, E., Sotak, T., Sparks, A., Stanley, H., Stone, H., Sutton, A., Svatek, A., Tabor, P., Tamerisa, A., Tamerisa, K., Tang, H., Tansey, J., Taylor, C., Taylor, T., Telford, B., Thomas, N., Thomas, S., Usmani, K., Vasquez, L., Vera, V., Villalon, D., Vinson, R., Wang, S., Ward-Moore, S., Warren, R., Washington, C., Watlington, S., Williams, G., Williamson, A., Wleczyk, R., Wooden, S., Worley, K., Wu, C., Wu, Y., Wu, Y. F., Zhou, J., Zorrilla, S., Nelson, D., Weinstein, G., and Gibbs, R.

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

2 (bases 1 to 158286)
 Direct Submission
 Worley, K.C.
 Submitted (19-FEB-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
 3 (bases 1 to 158286)
 Direct Submission
 Worley, K.C.
 Submitted (13-JUL-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
 On Jul 12, 2002 this sequence version replaced gi:18701102.

 Center: Baylor College of Medicine
 Center code: BCM
 Web site: <http://www.hgsc.bcm.tmc.edu/>
 Contact: hgsc-help@bcm.tmc.edu

 Project Information
 Center project name: GMBI
 Center clone name: CH230-240A17

 Summary Statistics

 Sequencing vector: Plasmid;
 Chemistry: Dye-terminator Big Dye; 100% of reads
 Assembly program: Phrap; version 0.990329
 Consensus quality: 103746 bases at least Q40
 Consensus quality: 108208 bases at least Q30
 Consensus quality: 111376 bases at least Q20

 * NOTE: Estimated insert size may differ from sequence length
 * (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 68 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.
 *
 * 1244: contig of 1244 bp in length
 * 1245 1344: gap of unknown length
 * 1345 2875: contig of 1531 bp in length
 * 2876 2975: gap of unknown length
 * 2976 4032: contig of 1057 bp in length
 * 4033 4132: gap of unknown length
 * 4133 5396: contig of 1264 bp in length
 * 5397 5496: gap of unknown length
 * 5497 5969: contig of 1473 bp in length
 * 5970 7069: gap of unknown length
 * 7070 8370: contig of 1301 bp in length
 * 8371 8470: gap of unknown length
 * 8471 9974: contig of 1504 bp in length
 * 9975 10074: gap of unknown length
 * 10075 11969: contig of 1895 bp in length
 * 11970 12069: gap of unknown length
 * 12070 13421: contig of 1352 bp in length
 *
 * 13422 13521: gap of unknown length
 * 13522 14649: contig of 1128 bp in length
 * 14650 14749: gap of unknown length
 * 14750 16351: contig of 1602 bp in length
 * 16352 16451: gap of unknown length
 * 16452 17940: contig of 1489 bp in length
 * 17941 18040: gap of unknown length
 * 18041 19287: contig of 1247 bp in length
 * 19288 19387: gap of unknown length
 * 19388 20768: contig of 1381 bp in length
 * 20769 20868: gap of unknown length
 * 20869 22395: contig of 1527 bp in length
 * 22396 22495: gap of unknown length
 * 22496 23636: contig of 1141 bp in length
 * 23637 23736: gap of unknown length
 * 23737 25287: contig of 1551 bp in length
 * 25288 25387: gap of unknown length
 * 25388 26570: contig of 1183 bp in length
 * 26571 26670: gap of unknown length
 * 26671 28078: contig of 1408 bp in length
 * 28079 28178: gap of unknown length
 * 28179 29573: contig of 1395 bp in length
 * 29574 29673: gap of unknown length
 * 29674 31324: contig of 1651 bp in length
 * 31325 31424: gap of unknown length
 * 31425 32599: contig of 1175 bp in length
 * 32600 32699: gap of unknown length
 * 32700 34751: contig of 2052 bp in length
 * 34752 34851: gap of unknown length
 * 34852 36020: contig of 1169 bp in length
 * 36021 36120: gap of unknown length
 * 36121 37941: contig of 1821 bp in length
 * 37942 38041: gap of unknown length
 * 38042 39144: contig of 1103 bp in length
 * 39145 39244: gap of unknown length
 * 39245 41215: contig of 1971 bp in length
 * 41216 41315: gap of unknown length
 * 41316 43354: contig of 2039 bp in length
 * 43355 43454: gap of unknown length
 * 43455 45504: contig of 2050 bp in length
 * 45505 45604: gap of unknown length
 * 45605 47070: contig of 1466 bp in length
 * 47071 47170: gap of unknown length
 * 47171 48306: contig of 1136 bp in length
 * 48307 48406: gap of unknown length
 * 48407 49467: contig of 1061 bp in length
 * 49468 49567: gap of unknown length
 * 49568 51036: contig of 1469 bp in length
 * 51037 51136: gap of unknown length
 * 51137 53712: contig of 2576 bp in length
 * 53713 53812: gap of unknown length
 * 53813 55162: contig of 1350 bp in length
 * 55163 55262: gap of unknown length
 * 55263 57096: contig of 1834 bp in length
 * 57097 57196: gap of unknown length
 * 57197 58767: contig of 1571 bp in length
 * 58768 58867: gap of unknown length
 * 58868 60411: contig of 1544 bp in length
 * 60412 60511: gap of unknown length
 * 60512 62118: contig of 1607 bp in length
 * 62119 62218: gap of unknown length
 * 62219 63943: contig of 1725 bp in length
 * 63944 64043: gap of unknown length
 * 64044 66201: contig of 2158 bp in length
 * 66202 66301: gap of unknown length
 * 66302 68867: contig of 2566 bp in length
 * 68868 68967: gap of unknown length
 * 68968 71412: contig of 2445 bp in length
 * 71413 71512: gap of unknown length
 * 71513 74169: contig of 2657 bp in length
 * 74170 74269: gap of unknown length
 * 74270 78013: contig of 3744 bp in length
 * 78014 78113: gap of unknown length

```

* 78114 80476: contig of 2363 bp in length
* 80477 80576: gap of unknown length
* 80577 83565: contig of 2989 bp in length
* 83566 83665: gap of unknown length
* 83666 83949: contig of 2284 bp in length
* 83949 86049: gap of unknown length
* 86050 89250: contig of 3201 bp in length
* 89251 89350: gap of unknown length
* 89351 91850: contig of 2500 bp in length
* 91851 91950: gap of unknown length
* 91951 93870: contig of 1920 bp in length
* 93871 93970: gap of unknown length
* 93971 96867: contig of 2897 bp in length
* 96868 96967: gap of unknown length
* 96968 97066: contig of 2739 bp in length
* 97067 97806: gap of unknown length
* 97807 101862: contig of 2056 bp in length

```

Query Match 14.4%; Score 19; DB 2; Length 158286;

Best Local Similarity 100.0%; Pred. No. 7.7;

Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 82 CTGATGACCAAGAGACCA 100

Db 76185 CTGATGACCAAGAGACCA 76203

RESULT 33

AC120092/c

LOCUS

DEFINITION

AC120092

AC120092

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

1 (bases 1 to 159642)

Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-Osman,F.R., Allen,C.,

Alsbrooks,S.L., Amarantunge,H.C., Are,J.R., Ayale,M., Banks,T.,

Barbora,J., Benton,J., Binage,K., Blankenburg,K., Bonnin,D.,

Bouck,J., Bowie,S., Brileva,M., Brown,E., Brown,M., Bryant,N.P.,

Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,

Carroll,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,

Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,

Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,

Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,

Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,

Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,

Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,

Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,

Gabis,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,

Garrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,

Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J.,

Hernandez,O., Hodgson,A., Hogue,M., Holloway,C., Hollins,B.,

Homs,J.F., Howard,S., Huber,J., Hulyk,S., Hume,J., Jackson,I.E.,

Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S.,

Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C.,

Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.,

Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W., Louised,H.,

Lozano,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,

Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E.,

Massey,E., Mawhiney,E., McLeod,M.P., Meador,M., Mei,G., Metzker,M.,

Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S.,

Moser,M., Neal,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N.,

Nguyen,N., Nickerson,E., Nwokenkwo,S., Ogih,M., Okwuonu,G.,

Oragunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L.,

Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y.,

Rives,M., Rojas,A., Rojibokan,I., Rolfe,M., Ruiz,S., Savery,G.,

Scherer,S., Scott,G., Shen,H., Shooshtari,N., Sisson,I.,

Sodergren,E., Sonaika,T., Sparks,A., Stanley,H., Stone,H.,
Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H.,
Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S.,
Usmani,K., Vasquez,L., Vera,V., Villalon,D., Vinson,R., Wang,Q.,
Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S.,
Williams,G., Williamson,A., Wleczkyk,R., Wooden,S., Worley,K.,
Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,
Weinstock,G. and Gibbs,R.
Direct Submission
Unpublished
2 (bases 1 to 159642)
Worley,K.C.
Direct Submission
Submitted (03-MAY-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 159642)
Worley,K.C.
Direct Submission
Submitted (14-JUL-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Jul 14, 2002 this sequence version replaced gi:21240251.

Center: Baylor College of Medicine

Center code: BCM

Web site: <http://www.hgsc.bcm.tmc.edu/>

Contact: hgsc-help@bcm.tmc.edu

----- Project Information

Center project name: GVZX

Center clone name: CH230-420B3

----- Summary Statistics

Sequencing vector: Plasmid

Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.990329

Consensus quality: 113642 bases at least Q40

Consensus quality: 118756 bases at least Q30

Consensus quality: 121702 bases at least Q20

* NOTE: Estimated insert size may differ from sequence length

(see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).

* NOTE: This is a 'working draft' sequence. It currently

* consists of 52 contigs. The true order of the pieces

* is not known and their order in this sequence record is

* arbitrary. Gaps between the contigs are represented as

* runs of N, but the exact sizes of the gaps are unknown.

* This record will be updated with the finished sequence

* as soon as it is available and the accession number will

* be preserved.

* 1 1114: contig of 1114 bp in length

* 1115 1214: gap of unknown length

* 1215 2253: contig of 1039 bp in length

* 2254 2353: gap of unknown length

* 2354 3777: contig of 1424 bp in length

* 3778 3877: gap of unknown length

* 3878 4982: contig of 1105 bp in length

* 4983 5082: gap of unknown length

* 5083 6267: gap of unknown length

* 6267 6366: gap of unknown length

* 6367 7571: contig of 1205 bp in length

* 7572 7671: gap of unknown length

* 7672 8822: contig of 1051 bp in length

* 8823 8822: gap of unknown length

* 8823 10076: contig of 1254 bp in length

* 10077 10176: gap of unknown length

* 10177 11560: contig of 1384 bp in length

* 11561 11660: gap of unknown length

* 11661 12728: contig of 1068 bp in length

* 12729 12828: gap of unknown length

* 12829 13916: contig of 1088 bp in length

* 13917 14016: gap of unknown length

* 14017 15947: contig of 1931 bp in length

* 15948 16047: gap of unknown length

* 16048 17736: contig of 1689 bp in length
* 17737 17836: gap of unknown length
* 17837 19231: contig of 1395 bp in length
* 19231 19331: gap of unknown length
* 19331 20481: contig of 1150 bp in length
* 20481 20581: gap of unknown length
* 20581 22886: contig of 2305 bp in length
* 22886 22987: gap of unknown length
* 22987 24533: contig of 1547 bp in length
* 24533 24634: gap of unknown length
* 24634 27351: contig of 2718 bp in length
* 27351 27451: gap of unknown length
* 27451 28985: contig of 1534 bp in length
* 28985 29086: gap of unknown length
* 29086 31338: contig of 2253 bp in length
* 31338 31438: gap of unknown length
* 31438 33141: contig of 1703 bp in length
* 33141 33241: gap of unknown length
* 33241 35675: contig of 2434 bp in length
* 35675 35775: gap of unknown length
* 35775 37592: contig of 1817 bp in length
* 37592 37693: gap of unknown length
* 37693 39329: contig of 1637 bp in length
* 39329 40854: gap of unknown length
* 40854 40754: gap of unknown length
* 40754 43168: contig of 2414 bp in length
* 43168 43269: gap of unknown length
* 43269 45414: contig of 2146 bp in length
* 45414 45515: gap of unknown length
* 45515 49320: contig of 3806 bp in length
* 49320 49420: gap of unknown length
* 49420 52472: contig of 3052 bp in length
* 52472 52572: gap of unknown length
* 52572 55255: contig of 2683 bp in length
* 55255 59303: gap of unknown length
* 59303 59403: gap of unknown length
* 59403 64041: gap of unknown length
* 64041 67011: contig of 2871 bp in length
* 67011 70301: gap of unknown length
* 70301 70401: gap of unknown length
* 70401 73883: contig of 3482 bp in length
* 73883 73983: gap of unknown length
* 73983 77036: contig of 3053 bp in length
* 77036 77136: gap of unknown length
* 77136 80975: contig of 3839 bp in length
* 80975 84129: contig of 3054 bp in length
* 84129 84230: gap of unknown length
* 84230 87620: contig of 3391 bp in length
* 87620 87721: gap of unknown length
* 87721 90509: contig of 2789 bp in length
* 90509 90609: gap of unknown length
* 90609 94032: contig of 3423 bp in length
* 94032 94132: gap of unknown length
* 94132 97299: contig of 3167 bp in length
* 97299 97399: gap of unknown length
* 97399 100648: contig of 3249 bp in length
* 100648 100748: gap of unknown length
* 100748 103849: contig of 3101 bp in length
* 103849 103850: gap of unknown length
* 103850 109523: contig of 5574 bp in length
* 109523 109624: gap of unknown length
* 109624 114061: contig of 4438 bp in length
* 114061 114161: gap of unknown length
* 114161 119832: contig of 5671 bp in length
* 119832 119932: gap of unknown length
* 119932 125547: contig of 5615 bp in length
* 125547 125648: gap of unknown length
* 125648 132259: contig of 6612 bp in length

* 132260 13259: gap of unknown length
* 13259 137599: contig of 5240 bp in length
* 137599 137600: gap of unknown length
* 137600 146877: contig of 9178 bp in length
* 146877 146977: gap of unknown length
* 146977 159642: contig of 12665 bp in length.
* 159642 Location/Qualifiers
* 1. .159642
* /organism="Rattus norvegicus"
* /db_xref="taxon:10116"

Query Match 14.4%; Score 19; DB 2; Length 159642;
Best Local Similarity 100.0%; Pred. No. 7.7;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 GGAAGAAATTCGACGCTTG 39
|||||
DB 44463 GGAAGAAATTCGACGCTTG 44445

RESULT 34
AC079549
LOCUS AC079549 160334 bp DNA linear HTG 02-SEP-2000
DEFINITION Mus musculus clone RP23-421E13, WORKING DRAFT SEQUENCE, 36
unordered pieces.
AC079549
AC079549.1 GI:9964914
VERSION HTG; HTGS_PHASE1; HTGS_DRAFT.
KEYWORDS Mus musculus.
SOURCE Mus musculus.
ORGANISM Mus musculus.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
1 (bases 1 to 160334)
TITLE Sequencing of Mouse
JOURNAL Unpublished
AUTHORS DOE Joint Genome Institute.
REFERENCE 2 (bases 1 to 160334)
TITLE Direct Submission
AUTHORS DOE Joint Genome Institute.
JOURNAL Submitted (02-SEP-2000) Production Sequencing Facility, DOE Joint
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
-----Genome Center
Center: Joint Genome Institute
Center Code: JGI
Web site: <http://www.jgi.doe.gov>

Project Information
Center Project Name: 1898645
Center clone name: RPCI-23_421E13

Summary Statistics
Consensus quality: 130033 bases at least Q40
Consensus quality: 148274 bases at least Q30
Consensus quality: 151667 bases at least Q20
Estimated insert size: 197000; agarose-fp estimation
Estimated insert size: 156834; sum-of-contigs estimation
Quality coverage: 3.52 in Q20 bases; agarose-fp estimation
Quality coverage: 4.42 in Q20 bases; sum-of-contigs estimation.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 36 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 1191: contig of 1191 bp in length
* 1192 1291: gap of unknown length
* 1292 2391: contig of 1100 bp in length
* 2392 2491: gap of unknown length
* 2492 3695: contig of 1204 bp in length
* 3696 3795: gap of unknown length
* 3796 4803: contig of 1008 bp in length

```
* 4804 4903: gap of unknown length.
* 4904 6037: contig of 1134 bp in length
* 6038 6137: gap of unknown length
* 6138 7603: contig of 1466 bp in length
* 7604 7703: gap of unknown length
* 7704 8945: contig of 1242 bp in length
* 8946 9045: gap of unknown length
* 9046 11077: contig of 2032 bp in length
* 11078 11177: gap of unknown length
* 11178 13117: contig of 1940 bp in length
* 13118 13217: gap of unknown length
* 13218 14463: contig of 1246 bp in length
* 14464 14563: gap of unknown length
* 14564 16339: contig of 1676 bp in length
* 16240 16339: gap of unknown length
* 16340 18236: contig of 1897 bp in length
* 18237 18336: gap of unknown length
* 18337 20463: contig of 2127 bp in length
* 20464 20563: gap of unknown length
* 20564 23184: contig of 2621 bp in length
* 23185 23284: gap of unknown length
* 23285 25190: contig of 1906 bp in length
* 25191 25290: gap of unknown length
* 25291 27703: contig of 2413 bp in length
* 27704 27803: gap of unknown length
* 27804 29855: contig of 2052 bp in length
* 29856 29955: gap of unknown length
* 29956 33001: contig of 3046 bp in length
* 33002 33101: gap of unknown length
* 33102 34938: contig of 1837 bp in length
* 34939 35038: gap of unknown length
* 35039 35668: contig of 3530 bp in length
* 35669 38668: gap of unknown length
* 38669 40977: contig of 2309 bp in length
* 40978 41077: gap of unknown length
* 41078 44724: contig of 3647 bp in length
* 44725 46888: gap of unknown length
* 46889 46988: contig of 2064 bp in length
* 46989 52032: gap of unknown length
* 52033 52132: contig of 5044 bp in length
* 52133 57669: contig of 5537 bp in length
* 57670 57769: gap of unknown length
* 57770 63675: contig of 5906 bp in length
* 63676 63775: gap of unknown length
* 63776 67199: contig of 3424 bp in length
* 67200 67299: gap of unknown length
* 67300 73134: contig of 5835 bp in length
* 73135 73233: gap of unknown length
* 73235 81223: contig of 7989 bp in length
* 81224 81323: gap of unknown length
* 81324 87283: contig of 5960 bp in length
* 87284 87383: gap of unknown length
* 87384 89401: contig of 2018 bp in length
* 89402 89501: gap of unknown length
* 89502 100676: contig of 11175 bp in length
* 100677 100776: gap of unknown length
* 100778 116153: contig of 15377 bp in length
* 116154 116253: gap of unknown length
* 116254 127943: contig of 11690 bp in length
* 127944 128043: gap of unknown length
* 128044 143038: contig of 14995 bp in length
* 143039 143138: gap of unknown length
* 143139 160334: contig of 17196 bp in length.
FEATURES
    source
    1. .160334
        /organism="Mus musculus"
        /db_xref="taxon:10090"
        /clone_lib="RPCI mouse BAC library 23"
        /clones="RP23-421E13"
45827 a 32566 c 33459 g 44842 t 3640 others
```

```
Query Match 14.4%; Score 19; DB 2; Length 160334;
Best Local Similarity 100.0%; Pred. No. 7.7;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 43 AAGGCCAAGCTGAAGGCCA 61
    |||||
Db 39137 AAGGCCAAGCTGAAGGCCA 39155

RESULT 35
AC034207 163942 bp DNA linear PRI 03-NOV-2001
LOCUS Homo sapiens chromosome 5 clone CTC-329H14, complete sequence.
DEFINITION AC034207
VERSION AC034207.6 GI:16603986
KEYWORDS HTG.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 163942)
AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.
TITLE Direct Submission
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 163942)
AUTHORS DOE Joint Genome Institute.
TITLE Direct Submission
JOURNAL Submitted (05-APR-2000) Production Sequencing Facility, DOE Joint
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
REFERENCE 3 (bases 1 to 163942)
AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.
TITLE Direct Submission
JOURNAL Submitted (30-AUG-2000) DOE Joint Genome Institute, 2800 Mitchell
Drive, Walnut Creek, CA 94598, USA
REFERENCE 4 (bases 1 to 163942)
AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.
TITLE Direct Submission
JOURNAL Submitted (03-OCT-2001) DOE Joint Genome Institute, 2800 Mitchell
Drive, Walnut Creek, CA 94598, USA
REFERENCE 5 (bases 1 to 163942)
AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.
TITLE Direct Submission
JOURNAL Submitted (03-NOV-2001) DOE Joint Genome Institute, 2800 Mitchell
Drive, Walnut Creek, CA 94598, USA
COMMENT On Nov 3, 2001 this sequence version replaced gi:15887323.
Draft Sequence Produced by DOE Joint Genome Institute
www.jgi.doe.gov
Finishing Completed at Stanford Human Genome Center
www.shgc.stanford.edu
Quality: Phrap Quality >=40 99% of Sequence;
Estimated Total Number of Errors is 1.2.
FEATURES
    source
    1. .163942
        /organism="Homo sapiens"
        /db_xref="taxon:9606"
        /chromosome="5"
        /clone="CTC-329H14"
BASE COUNT 52833 a 32579 c 31594 g 46936 t
ORIGIN
14.4%; Score 19; DB 9; Length 163942;
Best Local Similarity 100.0%; Pred. No. 7.7;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 90 CAAAGAGACACAGACGAG 108
    |||||
Db 66703 CAAAGAGACACAGACGAG 66721

RESULT 36
AC020685 166606 bp DNA linear HTG 08-APR-2000
LOCUS Homo sapiens chromosome 15 clone RP11-414J1 map 15, WORKING DRAFT
DEFINITION
```

ACCESSION VERSION KEYWORDS SOURCE ORGANISM

SEQUENCE, 14 unordered pieces.

AC020685
AC020685.3 GI:7528075
HTG: HTGS_PHASE1; HTGS_DRAFT.
Homo sapiens.
Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 166606)

Birren, B., Linton, L., Nusbaum, C. and Lander, E.

Homo sapiens chromosome 15, clone RP11-414J1

Unpublished

2 (bases 1 to 166606)

Birren, B., Linton, L., Nusbaum, C., Lander, E., Abraham, H., Allen, N.,
Anderson, S., Baldwin, J., Barna, N., Beckerly, R., Bada, F.,

Boguslavsky, L., Boukhalter, B., Brown, A., Burkett, G., Castle, A.,

Chapel, Y., Collangelo, M., Collins, S., Collymore, A., Cooke, P.,

DeAtellano, K., Dewar, K., Domino, M., Doyle, M., Fenestor, J.,

Ferreira, P., Fitzhugh, W., Forrest, C., Gage, D., Galagan, J.,

Gardyna, S., Grant, G., Hagos, B., Heaford, A., Horton, L.,

Howland, J.C., Johnson, R., Jones, C., Kann, L., Karatas, A., Klein, J.,

Landers, T., Lehoczy, J., Levine, R., Liu, C., Liu, G., Locke, K.,

Macdonald, P., Marquis, N., McEwan, P., McGurk, A., McKernan, K.,

McPheeters, R., Meldrum, J., Meneus, L., Morrow, J., Naylor, J.,

Norman, C.H., O'Connor, T., O'Donnell, P., Oliver, T.M., Peterson, K.,

Pierre, N., Pisani, C., Pollara, V., Raymond, C., Riley, R., Rothman, D.,

Roy, A., Santos, R., Severy, P., Spencer, B., Stange-Thomann, N.,

Stojanovic, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J.,

Tirrell, A., Vassiliev, H., Viel, R., Vo, A., Wu, X., Wyman, D., Ye, W.J.,

Zimmer, A. and Zody, M.

Direct Submission

Submitted (08-JAN-2000) Whitehead Institute/MIT Center for Genome

Research, 320 Charles Street, Cambridge, MA 02141, USA

On Apr 8, 2000 this sequence version replaced gi:6751753.

All repeats were identified using RepeatMasker:

Smit, A.F.A. & Green, P. (1996-1997)

http://ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WTBPR

Web site: http://www-seq.wi.mit.edu

Contact: sequence_submissions@genome.wi.mit.edu

----- Project Information

Center project name: L4965

Center clone name: 414_J_1

----- Summary Statistics

Sequencing vector: M13; M7815; 100% of reads

Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.960731

Consensus quality: 159639 bases at least Q40

Consensus quality: 162851 bases at least Q30

Consensus quality: 163954 bases at least Q20

Insert size: 164000; agarose-fp

Insert size: 165306; sum-of-contigs

Quality coverage: 5.6 in Q20 bases; agarose-fp

Quality coverage: 5.5 in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently
* consists of 14 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

* 1 1059: contig of 1059 bp in length
* 1060 1159: gap of 100 bp
* 1160 6470: contig of 5311 bp in length
* 6471 6570: gap of 100 bp
* 6571 9999: contig of 3429 bp in length
* 10000 10099: gap of 100 bp
* 10100 14075: contig of 3976 bp in length
* 14076 14175: gap of 100 bp

* 14176 19098: contig of 4923 bp in length
* 19099 19198: gap of 100 bp
* 19199 26231: contig of 7033 bp in length
* 26232 26331: gap of 100 bp
* 26332 35257: contig of 8926 bp in length
* 35258 35357: gap of 100 bp
* 35358 49626: contig of 14269 bp in length
* 49627 49726: gap of 100 bp
* 49727 64090: contig of 14364 bp in length
* 64091 64190: gap of 100 bp
* 64191 81200: contig of 17010 bp in length
* 81201 81300: gap of 100 bp
* 81301 97396: contig of 16096 bp in length
* 97397 97496: gap of 100 bp
* 97497 115927: contig of 18431 bp in length
* 115928 116027: gap of 100 bp
* 116028 136724: contig of 20697 bp in length
* 136725 136824: gap of 100 bp
* 136825 166606: contig of 29782 bp in length.

FEATURES

source

1. 166606
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="15"
/map="15"

/clone="RP11-414J1"

/clone_lib="RPCI-11 Human Male BAC"

1. 1059

/note="assembly_fragment"

misc_feature 1160..6470

/note="assembly_fragment"

misc_feature 6571..9999

/note="assembly_fragment"

misc_feature 10100..14075

/note="assembly_fragment"

misc_feature 14176..19098

/note="assembly_fragment"

misc_feature 19199..26231

/note="assembly_fragment"

misc_feature 26332..35257

/note="assembly_fragment"

misc_feature 35358..49626

/note="assembly_fragment"

clone_end:SP6

vector_side:right

49727..64090

/note="assembly_fragment"

misc_feature 64191..81200

/note="assembly_fragment"

misc_feature 81301..97396

/note="assembly_fragment"

clone_end:T7

vector_side:left

97497..115927

/note="assembly_fragment"

misc_feature 116028..136724

/note="assembly_fragment"

misc_feature 136825..166606

/note="assembly_fragment"

BASE COUNT 48221 a 33530 c 33827 g 49727 t 1301 others

ORIGIN

Query Match 14.4%; Score 19; DB 2; Length 166606;
Best Local Similarity 100.0%; Pred. No. 7.7;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 57 GGCCACAGAGATGCAGAAG 75

|||||

Db 70573 GGCCACAGAGATGCAGAAG 70591

RESULT 37

AC126443/c

```

LOCUS      AC126443               168766 bp    DNA    linear    HTG 06-JUL-2002
DEFINITION Mus musculus chromosome UNK clone RP23-464J8, WORKING DRAFT
SEQUENCE, 17 unordered pieces.
ACCESSION  AC126443
VERSION    AC126443.1  GI:21699723
KEYWORDS   HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE     house mouse.
ORGANISM   Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE  1 (bases 1 to 168766)
AUTHORS   McPherson,J.D. and Waterston,R.H.
TITLE     The sequence of Mus musculus clone
JOURNAL   Unpublished
REFERENCE  2 (bases 1 to 168766)
AUTHORS   McPherson,J.D. and Waterston,R.H.
TITLE     Direct Submission
JOURNAL   Direct Submission
REFERENCE  3 (bases 1 to 168766)
AUTHORS   Submitted (06-JUL-2002) Genome Sequencing Center, 4444 Forest Park
TITLE     Parkway, St. Louis, MO 63108, USA
JOURNAL   Submitted (06-JUL-2002) Genome Sequencing Center, 4444 Forest Park
          Parkway, St. Louis, MO 63108, USA

COMMENT   ----- Genome Center
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc/index.shtml
Contact: submissions@watson.wustl.edu
          ----- Project Information -----
Center project name: M_BA0464J08
          ----- Summary Statistics -----
Sequencing vector: M13; 0%
Chemistry: Dye-primer ET; 0% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 165197 bases at least Q40
Consensus quality: 166166 bases at least Q30
Consensus quality: 166509 bases at least Q20
Insert size: 186000; agarose-fp
Insert size: 168200; sum-of-contigs
Quality coverage: 0.00 in Q20 bases; agarose-fp
Quality coverage: 10.38 in Q20 bases; sum-of-contigs
          -----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 17 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1 1585: contig of 1585 bp in length
* 1586 1685: gap of unknown length
* 1686 3555: contig of 1870 bp in length
* 3556 6530: gap of unknown length
* 6531 10605: contig of 2875 bp in length
* 10606 10605: gap of unknown length
* 10606 10705: contig of 3975 bp in length
* 10706 15789: contig of 5084 bp in length
* 15790 15889: gap of unknown length
* 15890 22248: contig of 6359 bp in length
* 22249 22348: gap of unknown length
* 22349 29057: contig of 6709 bp in length
* 29058 29158: gap of unknown length
* 29158 37663: contig of 8406 bp in length
* 37664 43022: gap of unknown length
* 43023 43122: contig of 5359 bp in length
* 43123 48872: contig of 5750 bp in length
* 48873 48972: gap of unknown length
* 48973 67906: gap of unknown length
* 67907 68006: gap of unknown length
* 68007 79408: contig of 11301 bp in length
* 79409 103114: contig of 23707 bp in length
* 103115 116372: gap of unknown length
* 116373 116372: contig of 13058 bp in length
* 116373 134688: gap of unknown length
* 134689 134788: contig of 18934 bp in length
* 134789 158534: gap of unknown length
* 158535 158635: gap of unknown length
* 158636 168766: contig of 10132 bp in length.

FEATURES             Location/Qualifiers
     source            1..168766
                        /organism="Mus musculus"
                        /db_xref="taxon:10090"
                        /chromosome="UNK"
                        /clone="RP23-464J8"
     misc_feature       1..1585
                        /note="assembly_name:Contig10"
     misc_feature       1686..3555
                        /note="assembly_name:Contig11"
     misc_feature       3556..6530
                        /note="assembly_name:Contig12"
     misc_feature       6531..10605
                        /note="assembly_name:Contig13"
     misc_feature       10706..15789
                        /note="assembly_name:Contig14"
     misc_feature       15890..22248
                        /note="assembly_name:Contig15"
     misc_feature       22349..29057
                        /note="assembly_name:Contig16"
     misc_feature       29158..37563
                        /note="assembly_name:Contig17"
     misc_feature       37664..43022
                        /note="assembly_name:Contig18"
     misc_feature       43123..48872
                        /note="assembly_name:Contig19"
     misc_feature       48973..67906
                        /note="assembly_name:Contig20"
     misc_feature       68007..79307
                        /note="assembly_name:Contig21"
     misc_feature       79408..103114
                        /note="assembly_name:Contig22"
     misc_feature       103215..116272
                        /note="assembly_name:Contig23"
     misc_feature       116373..134687
                        /note="assembly_name:Contig24"
     misc_feature       134788..158534
                        /note="assembly_name:Contig25"
     misc_feature       158635..168766
                        /note="assembly_name:Contig26"

BASE COUNT  52459 a 32388 c 33335 g 48916 t 1668 others
ORIGIN
Query Match      14.4%; Score 19; DB 2; Length 168766;
Best Local Similarity 100.0%; Pred. No. 7.7;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      43 AAGGCCAAGCTGAAGGCCA 61
        |||||
Db      109635 AAGGCCAAGCTGAAGGCCA 109617

RESULT 38
AL845505/c      172191 bp    DNA    linear    HTG 16-AUG-2002
LOCUS          Mus musculus chromosome X clone RP23-115D12, *** SEQUENCING IN
DEFINITION    PROGRESS ***, 21 unordered pieces.
ACCESSION     AL845505

```


/note="assembly_fragment:01013

fragment_chain:4

clone_end:r7

vector_side:right"

BASE COUNT 49464 a 32131 c 32707 g 55878 t 2011 others

ORIGIN

Query Match 14.4%; Score 19; DB 2; Length 172191;
Best Local Similarity 100.0%; Pred. No. 7.7;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 43 AAGGCCAAGCTGAAGGCCA 61

Db 79245 AAGGCCAAGCTGAAGGCCA 79227

RESULT 39

AC122629

LOCUS

DEFINITION

***, 67 unordered pieces.

ACCESSION

AC122629.2 GI:21731433

VERSION

HTG; HTGS-PHASE1.

KEYWORDS

SOURCE

ORGANISM

Rattus norvegicus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;

Rattus.

1 (bases 1 to 172797)

Muzny D.M., Adams C., Adio-Oduola B., Ali-Osman F.R., Allen C.,

Albrooks S.L., Amarantunge H.C., Are J.R., Ayala M., Banks T.,

Barbata J., Benton J., Blincke K., Blankenburg K., Bonnin D.,

Bouck J., Bowler S., Brice M., Brown E., Brown M., Bryant N.P.,

Buhal C., Burch P., Burkett C., Burrell K.L., Byrd N.C.,

Carron T.F., Carter M., Cavazos S.R., Chacko J., Chavez D.,

Chen G., Chen R., Chen Z., Chowdhry I., Christopoulos C.,

Cleveland C.D., Cox C., Coyle M.D., Dathorne S.R., David R.,

Davila M.L., Davis C., Davy-Carroll L., Dederich D.A.,

Delaney K.R., Delgado O., Denn A.L., Ding Y., Dinh H.H.,

Douthwaite K.J., Draper H., Dugan-Rocha S., Durbin K.J.,

Earnhart C., Edgar D., Edwards C.C., Elhaj C., Escotto M.,

Falls T., Ferraguto D., Flagg N., Ford J., Foster P., Frantz P.,

Gabis A., Gao J., Garcia A., Garner T., Garza N., Gill R.,

Gorrell J.H., Guevara W., Gunaratne P., Hale S., Hamilton K.,

Harris C., Harris K., Hart M., Havlak P., Hawes A., Hernandez J.,

Hernandez O., Hodgson A., Hognes M., Holloway C., Hollins B.,

Homsif F., Howard S., Huber J., Hulyk S., Hume J., Jackson L.E.,

Jacobson B., Jia Y., Johnson R., Jolivet S., Joudah S.,

Karlsson E., Kelly S., Khan U., King L., Korvah J., Kovar C.,

Kratovic J., Kureshi A., Landry N., Leal B., Lewis L.C., Lewis L.,

Li J., Li Z., Lichtarge O., Lieu C., Liu J., Liu W., Louisegeed H.,

Lozano R.J., Lu X., Lucier A., Lucier R., Luna R., Ma J.,

Maheshwari M., Mapua P., Martin R., Martindale A., Martinez E.,

Massey E., Mahoney E., McLeod M.P., Meador M., Mei G., Metzker M.,

Miner G., Miner Z., Mitchell T., Mohabbat K., Morgan M., Morris S.,

Moser M., Neal D., Newton J., Newton N., Nguyen A., Nguyen N.,

Nguyen N., Nickerson E., Nwokenwo S., Ogih M., Okunolu G.,

Oragunye N., Oviedo R., Pace A., Payton B., Peery J., Perez L.,

Peters L., Picken R., Primus E., Pu L.L., Quiles M., Ren Y.,

Rives M., Rojas A., Rojubokan I., Rolfe M., Ruiz S., Savary G.,

Scherer S., Scott G., Shen H., Shoohtari N., Sisson I.,

Sodergren E., Sonaika T., Sparks A., Stanley H., Stone H.,

Stanton A., Svatek A., Tabor P., Tamerisa A., Tamerisa K., Tang H.,

Tansey J., Taylor C., Taylor T., Telford B., Thomas N., Thomas S.,

Usmani K., Vasquez L., Vera V., Villalón D., Vinson R., Wang Q.,

Wang S., Ward-Moore S., Warren R., Washington C., Watlington S.,

Williams G., Williamson A., Wleczyk R., Wooden S., Worley K.,

Wu C., Wu Y., Wu Y.F., Zhou J., Zorrilla S., Nelson D.,

Weinstock G. and Gibbs R.

Direct Submission

Unpublished

2 (bases 1 to 172797)

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Worley K.C.

Direct Submission

Submitted (25-MAY-2002) Human Genome Sequencing Center, Department

of Molecular and Human Genetics, Baylor College of Medicine, One

Baylor Plaza, Houston, TX 77030, USA

3 (bases 1 to 172797)

Worley K.C.

Direct Submission

Submitted (14-JUL-2002) Human Genome Sequencing Center, Department

of Molecular and Human Genetics, Baylor College of Medicine, One

Baylor Plaza, Houston, TX 77030, USA

On Jul 11, 2002 this sequence version replaced gi:21205946.

----- Genome Center

Center: Baylor College of Medicine

Center code: BCM

Web site: <http://www.hgsc.bcm.tmc.edu/>

Contact: hgsc-help@bcm.tmc.edu

----- Project Information

Center project name: GKCM

Center clone name: CH230-102A4

----- Summary Statistics

Sequencing vector: Plasmid;

Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.990329

Consensus quality: 113756 bases at least Q40

Consensus quality: 118828 bases at least Q30

Consensus quality: 123024 bases at least Q20

* NOTE: Estimated insert size may differ from sequence length

* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).

* NOTE: This is a 'working draft' sequence. It currently

* consists of 67 contigs. The true order of the pieces

* is not known and their order in this sequence record is

* arbitrary. Gaps between the contigs are represented as

* runs of N, but the exact sizes of the gaps are unknown.

* This record will be updated with the finished sequence

* as soon as it is available and the accession number will

* be preserved.

* 1 1758: contig of 1758 bp in length

* 1759 1858: gap of unknown length

* 1859 2964: contig of 1106 bp in length

* 2965 3064: gap of unknown length

* 3065 4560: contig of 1496 bp in length

* 4561 4660: gap of unknown length

* 4661 5819: contig of 1159 bp in length

* 5820 7366: contig of unknown length

* 7367 7466: contig of 1447 bp in length

* 7467 8958: gap of unknown length

* 8959 9058: contig of 1492 bp in length

* 9059 10424: contig of 1366 bp in length

* 10425 12064: contig of unknown length

* 12065 12164: gap of unknown length

* 12165 13615: contig of 1451 bp in length

* 13616 13715: gap of unknown length

* 13716 15101: contig of 1386 bp in length

* 15102 16254: gap of unknown length

* 16255 16354: gap of unknown length

* 16355 17443: contig of 1089 bp in length

* 17444 17543: gap of unknown length

* 17544 18862: contig of 1319 bp in length

* 18863 18962: gap of unknown length

* 18963 20482: contig of 1520 bp in length

* 20483 20582: gap of unknown length

* 20583 22101: contig of 1519 bp in length

* 22102 22201: gap of unknown length

* 22202 23489: contig of 1288 bp in length

* 23490 23589: gap of unknown length

* 23590 25078: contig of 1489 bp in length

* 25079 25178: gap of unknown length

* 25179 26268: contig of 1090 bp in length

* 26269 26268: contig of 1090 bp in length

* 26269 26268: contig of 1090 bp in length

* 26269 26268: contig of 1090 bp in length

* 26269 26268: contig of 1090 bp in length


```
*
26269 26368: gap of unknown length
26369 27550: contig of 1182 bp in length
27551 27650: gap of unknown length
27651 29238: contig of 1588 bp in length
29239 29338: gap of unknown length
29339 30567: contig of 1229 bp in length
30568 30667: gap of unknown length
30668 31946: contig of 1279 bp in length
31947 32046: gap of unknown length
32047 33276: contig of 1230 bp in length
33277 33376: gap of unknown length
33377 34712: contig of 1336 bp in length
34713 34812: gap of unknown length
34813 36059: contig of 1247 bp in length
36060 36159: gap of unknown length
36160 38183: contig of 2024 bp in length
38184 38283: gap of unknown length
38284 40179: contig of 1896 bp in length
40180 40279: gap of unknown length
40280 42485: contig of 2206 bp in length
42486 42585: gap of unknown length
42586 44371: contig of 1786 bp in length
44372 44471: gap of unknown length
44472 46529: contig of 2158 bp in length
46530 46729: gap of unknown length
46730 48613: contig of 1884 bp in length
48614 48713: gap of unknown length
48714 50116: contig of 1403 bp in length
50117 50216: gap of unknown length
50217 52877: contig of 2661 bp in length
52878 52977: gap of unknown length
52978 54957: contig of 1980 bp in length
54958 55057: gap of unknown length
55058 57019: contig of 1962 bp in length
57020 57119: gap of unknown length
57120 59360: contig of 2241 bp in length
59361 59460: gap of unknown length
59461 61172: contig of 1712 bp in length
61173 61272: gap of unknown length
61273 62405: contig of 1133 bp in length
62406 62505: gap of unknown length
62506 64814: contig of 2309 bp in length
64815 64915: gap of unknown length
64916 67708: contig of 2794 bp in length
67709 69632: contig of 1824 bp in length
69633 69732: gap of unknown length
69733 72062: contig of 2330 bp in length
72063 72162: gap of unknown length
72163 73921: contig of 1759 bp in length
73922 74021: gap of unknown length
74022 76561: contig of 2540 bp in length
76562 76661: gap of unknown length
76662 79481: contig of 2820 bp in length
79482 79581: gap of unknown length
79582 83000: contig of 3419 bp in length
83001 83100: gap of unknown length
83101 87071: contig of 3971 bp in length
87072 87171: gap of unknown length
87172 89546: contig of 2375 bp in length
89547 89647: gap of unknown length
89648 92393: contig of 2747 bp in length
92394 92493: gap of unknown length
92494 96986: contig of 4493 bp in length
96987 97086: gap of unknown length
97087 101666: contig of 4580 bp in length
101667 101766: gap of unknown length
101767 105032: contig of 3266 bp in length
105033 105132: gap of unknown length
105133 107532: contig of 2400 bp in length
107533 112074: contig of 4442 bp in length
112074: contig of 4442 bp in length
```

```
Query Match 14.4%; Score 19; DB 2; Length 172737;
Best Local Similarity 100.0%; Pred. No. 7.7;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 GGAGAAATGCCAGCTTG 39
    |||||
Db 112772 GGAAGAAATGCCAGCTTG 112790
    |||||

RESULT 40
AC008871/c 174138 bp DNA linear PRI 06-DEC-2001
LOCUS Homo sapiens chromosome 5 clone CTD-2197111, complete sequence.
DEFINITION AC008871
ACCESSION AC008871
VERSION AC008871.5 GI:17386232
KEYWORDS HTG.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 174138)
AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.
TITLE Direct Submission
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 174138)
AUTHORS DOE Joint Genome Institute.
TITLE Direct Submission
JOURNAL Submitted (03-AUG-1999) Production Sequencing Facility, DOE Joint
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
REFERENCE 3 (bases 1 to 174138)
AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.
TITLE Direct Submission
JOURNAL Submitted (06-DEC-2001) DOE Joint Genome Institute, 2800 Mitchell
Drive, Walnut Creek, CA 94598, USA
COMMENT On Dec 6, 2001 this sequence version replaced gi:9256065.
Draft Sequence Produced by DOE Joint Genome Institute
www.jgi.doe.gov
Finishing Completed at Stanford Human Genome Center
www.shgc.stanford.edu
Quality: Phrap Quality >=40 99.9% of Sequence;
Estimated Total Number of Errors is 0.1.

FEATURES
    source
    1..174138
        /organism="Homo sapiens"
        /db_xref="taxon:9606"
        /chromosome="5"
        /clone="CTD-2197111"

BASE COUNT 49893 a 31309 c 33213 g 59723 t
ORIGIN

Query Match 14.4%; Score 19; DB 9; Length 174138;
Best Local Similarity 100.0%; Pred. No. 7.7;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 90 CAAGAGACCACAGAGCAG 108
    |||||
Db 35324 CAAGAGACCACAGAGCAG 35306
    |||||

RESULT 41
AC112933 174670 bp DNA linear HTG 21-AUG-2002
LOCUS Mus musculus clone RP24-418J2, WORKING DRAFT SEQUENCE, 8 ordered
pieces
DEFINITION AC112933
ACCESSION AC112933
VERSION AC112933.4 GI:22381912
KEYWORDS HTG; HTGS_PHASE2; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 174670)
AUTHORS Birren,B., Nusbaum,C. and Lander,E.
```

TITLE JOURNAL REFERENCE AUTHORS

Mus musculus, clone RP24-418J2
Unpublished
2 (bases 1 to 174670)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
Anderson,S., Barna,N., Bastien,V., Boguslavskiy,L., Boukhgalter,B.,
Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B.,
Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cook,A.,
Cooke,P., DeArillano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S.,
Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,
Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
Kamat,A., Karatas,A., Kells,C., LaRocque,K., Lamazares,R.,
Landers,T., Lehoczy,J., Levine,R., Liu,G., MacLean,C.,
Macdonald,P., Major,J., Marquis,N., Matthews,C., McCarthy,M.,
McWana,V., McKernan,K., Meldrim,J., Meneus,L., Mihova,T.,
Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C.,
Noren,C.H., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J.,
Peterson,K., Phunkhang,P., Pierre,R., Rise,C., Rogov,P., Raymond,C.,
Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J.,
Rosetti,M., Roy,A., Santos,R., Schauer,S., Schuback,R., Seaman,S.,
Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
Topham,K., Travers,M., Travis,N., Triggilio,J., Vassiliev,H.,
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
Direct Submission
Submitted (25-FEB-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 174670)
Birren,B., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S.,
Barna,N., Bastien,V., Bloom,T., Boguslavskiy,L., Boukhgalter,B.,
Camarata,J., Chang,J., Chazaro,B., Choepel,Y., Collymore,A.,
Cook,A., Cooke,P., DeArillano,K., Dewar,K., Diaz,J.S., Dodge,S.,
Faro,S., Ferreira,P., FitzGerald,M., Gage,D., Galagan,J.,
Gardyna,S., Gord,S., Graham,L., Grand-Pierre,N., Hagos,B.,
Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A.,
Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K.,
Liu,G., MacLean,C., Macdonald,P., Major,J., Matthews,C.,
McCarthy,M., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V.,
Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H.,
O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K.,
Phunkhang,P., Pierre,R., Raymond,C., Retta,R., Rise,C., Rogov,P.,
Roman,J., Roy,A., Schauer,S., Schuback,R., Seaman,S., Severy,P.,
Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Talamas,J.,
Tesfaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H.,
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J.,
Zembek,L., Zimmer,A. and Zody,M.
Direct Submission

TITLE JOURNAL COMMENT

Submitted (21-AUG-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Aug 21, 2002 this sequence version replaced gi:21591839.
All repeats were identified using RepeatMasker:
Smit, A.P.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: <http://www-seq.wi.mit.edu>
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L20144
Center clone name: 418_J_2
----- Summary Statistics
Sequencing vector: Plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 172471 bases at least Q40
Consensus quality: 173097 bases at least Q30
Consensus quality: 173390 bases at least Q20
Insert size: 163000; agarose-fp
Quality coverage: 11.6 in Q20 bases; sum-of-contigs
Quality coverage: 10.9 in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently
* consists of 8 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
* 1 1991: contig of 1991 bp in length
* 1992 2091: gap of 100 bp
* 2092 4608: contig of 2517 bp in length
* 4609 4708: gap of 100 bp
* 4709 10380: contig of 5672 bp in length
* 10381 10480: gap of 100 bp
* 10481 19423: contig of 8943 bp in length
* 19424 19523: gap of 100 bp
* 19524 47741: contig of 28218 bp in length
* 47742 47841: gap of 100 bp
* 47842 76804: contig of 28963 bp in length
* 76805 76904: gap of 100 bp
* 76905 122708: contig of 45804 bp in length
* 122709 122808: gap of 100 bp
* 122809 174670: contig of 51862 bp in length.
----- Location/Qualifiers
1. 174670
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="RP24-418J2"
/clone_lib="RPCI-24 Male Mouse BAC"
1. 1991
/note="assembly_fragment"
2092. 4608
/note="assembly_fragment"
4709. 10380
/note="assembly_fragment"
10481. 19423
/note="assembly_fragment"
19524. 47741
/note="assembly_fragment"
47842. 76804
/note="assembly_fragment"
76905. 122708
/note="assembly_fragment"
122809. 174670
/note="assembly_fragment"
BASE COUNT 50336 a 37369 c 36806 g 49456 t
ORIGIN

Query Match 14.4%; Score 19; DB 2; Length 174670;
Best Local Similarity 100.0%; Pred. No. 7.7;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 43 AAGCCCAAGCTGAAGCCA 61
|||||
Db 40130 AAGCCCAAGCTGAAGCCA 40148

RESULT 42 AC118879

LOCUS AC118879 175037 bp DNA linear HTG 14-JUL-2002
DEFINITION Rattus norvegicus clone CH230-472K2, *** SEQUENCING IN PROGRESS
***, 58 unordered pieces.

ACCESSION AC118879
VERSION AC118879.2 GI:21747356
KEYWORDS HTG; HTGS; PHASE1.

SOURCE Norway rat.
ORGANISM Rattus norvegicus

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.

REFERENCE 1 (bases 1 to 175037)

AUTHORS

Muzny, D.M., Adams, C., Adio-Oduola, B., Ali-osman, F.R., Allen, C., Albrooks, S.L., Amaratunga, H.C., Are, J.R., Ayele, M., Banks, T., Barbra, J., Benton, J., Bimage, K., Blankenburg, K., Bonnin, D., Bouck, J., Brieva, M., Brown, E., Brown, M., Bryant, N.P., Buhay, C., Burch, P., Burkett, C., Burrell, K.L., Byrd, N.C., Caron, T.F., Carter, M., Cavazos, S.R., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Z., Chowdhry, I., Christopoulos, C., Cleveland, C.D., Cox, C., Coyle, M.D., Dathorne, S.R., David, R., Davila, M.L., Davis, C., Davy-Carroll, L., Dederich, D.A., Delaney, K.R., Delgado, O., Denn, A.L., Ding, Y., Dinh, H.H., Douthwaite, K.J., Draper, H., Dugan-Rocha, S., Durbin, K.J., Earnhart, C., Edgar, D., Edwards, C.C., Elhaj, C., Escotto, M., Falls, T., Ferraguto, D., Flagg, N., Ford, J., Foster, P., Frantz, P., Gabisi, A., Gao, J., Garcia, A., Garner, T., Garza, N., Gill, R., Gorrell, J.H., Guevara, W., Gunaratne, P., Hale, S., Hamilton, K., Harris, C., Harris, K., Hart, M., Havlak, P., Hawes, A., Hernandez, J., Hernandez, O., Hodgson, A., Hogues, M., Holloway, C., Hollins, B., Honsi, F., Howard, S., Huber, J., Hulyk, S., Hume, J., Jackson, L.E., Jacobson, B., Jia, F., Johnson, R., Jolivet, S., Joudah, S., Karlsson, E., Kelly, S., Khan, U., King, L., Korvah, J., Kovar, C., Kratovic, J., Kureshi, A., Landry, N., Leal, B., Lewis, L.C., Lewis, L., Li, J., Li, Z., Lichtarge, O., Lieu, C., Liu, J., Liu, W., Loulseg, H., Lozado, R.J., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J., Maheshwari, M., Mapua, P., Martin, R., Martindale, A., Martinez, E., Massey, E., Mawhiney, E., McLeod, M.P., Meador, M., Mei, G., Metzker, M., Miner, G., Miner, Z., Mitchell, T., Mohabbat, K., Morgan, M., Morris, S., Moser, M., Neal, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N., Nguyen, N., Nickerson, E., Nwokenkwo, S., Ogih, M., Okwuonu, G., Oraqunye, N., Oviado, R., Pace, A., Payton, B., Peery, J., Perez, L., Peters, L., Pickens, R., Primus, E., Pu, L.L., Quiles, M., Ren, Y., Rivers, M., Rojas, A., Rojibokan, I., Rolfe, M., Ruiz, S., Savery, G., Scherer, S., Scott, G., Shen, H., Shoohtari, N., Sisson, I., Sodergren, E., Sonaik, T., Sparks, A., Stanley, H., Stone, H., Sutton, A., Svatek, A., Tabor, P., Tamerisa, A., Tamerisa, K., Tang, H., Tansey, J., Taylor, C., Taylor, T., Telford, B., Thomas, N., Thomas, S., Usmani, K., Vaquez, L., Vera, Y., Villalon, D., Vinson, R., Wang, Q., Wang, S., Ward-Moore, S., Warren, R., Washington, C., Watlington, S., Williams, G., Williamson, A., Wleczyk, R., Wooden, S., Worley, K., Wu, C., Wu, Y., Wu, Y.F., Zhou, J., Zorrilla, S., Nelson, D., Weinstock, G. and Gibbs, R.

TITLE
JOURNAL
REFERENCE
2 (bases 1 to 175037)
Worley, K.C.

AUTHORS

Submitted (21-APR-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

3 (bases 1 to 175037)
Worley, K.C.

AUTHORS

Submitted (14-JUL-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

COMMENT

On Jul 14, 2002 this sequence version replaced gi:20258343.

Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
----- Project information
Center project name: GVYO
Center clone name: CH230-472K2
----- Summary Statistics
Sequencing vector: Plasmid:

Chemistry: Dye-terminator Big Dye: 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 138521 bases at least Q40
Consensus quality: 142923 bases at least Q30
Consensus quality: 146010 bases at least Q20

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).

* NOTE: This is a 'working draft' sequence. It currently consists of 58 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

1 1007: contig of 1007 bp in length
1008 1107: gap of unknown length
1108 2149: contig of 1042 bp in length
2150 2249: gap of unknown length
2250 3370: contig of 1121 bp in length
3371 3470: gap of unknown length
3471 4646: contig of 1176 bp in length
4647 4746: gap of unknown length
4747 5876: contig of 1130 bp in length
5877 5976: gap of unknown length
5977 7629: contig of 1553 bp in length
7530 9006: contig of 1377 bp in length
9007 9106: gap of unknown length
9107 10780: contig of 1674 bp in length
10781 10880: gap of unknown length
10881 12444: contig of 1564 bp in length
12445 12544: gap of unknown length
12545 14109: contig of 1565 bp in length
14110 14209: gap of unknown length
14210 15653: contig of 1444 bp in length
15654 15753: gap of unknown length
15754 16877: contig of 1024 bp in length
16878 19235: contig of 2358 bp in length
19236 19335: gap of unknown length
19336 20455: contig of 1120 bp in length
20456 21738: contig of 1183 bp in length
21739 21838: gap of unknown length
21839 23653: contig of 1815 bp in length
23654 23753: gap of unknown length
23754 25302: contig of 1549 bp in length
25303 25402: gap of unknown length
25403 26614: contig of 1212 bp in length
26615 26714: gap of unknown length
26715 28514: contig of 1800 bp in length
28515 28614: gap of unknown length
28615 30443: contig of 1829 bp in length
30444 30543: gap of unknown length
30544 32375: contig of 1832 bp in length
32376 32476: gap of unknown length
32477 34038: contig of 1562 bp in length
34039 34137: gap of unknown length
34138 35998: contig of 1861 bp in length
35999 36098: gap of unknown length
36099 37997: contig of 1898 bp in length
37998 40345: contig of 2249 bp in length
40346 40445: gap of unknown length
40446 42372: contig of 1927 bp in length
42373 42472: gap of unknown length
42473 45396: contig of 2924 bp in length
45397 45496: gap of unknown length
45497 48748: contig of 3252 bp in length
48749 48848: gap of unknown length
48849 51292: contig of 2444 bp in length
51293 51392: gap of unknown length
51393 54376: contig of 2984 bp in length
54377 54476: gap of unknown length
54477 57758: contig of 3282 bp in length
57759 57858: gap of unknown length
57859 60723: contig of 2865 bp in length
60724 60823: gap of unknown length
60824 63425: contig of 2602 bp in length

```

* 63426 63525: gap of unknown length
* 63526 66201: contig of 2676 bp in length
* 66202 66301: gap of unknown length
* 66302 69653: contig of 3352 bp in length
* 69654 72925: gap of unknown length
* 72926 73025: contig of 3172 bp in length
* 73026 75111: contig of 2386 bp in length
* 75112 75511: gap of unknown length
* 75512 78885: contig of 3374 bp in length
* 78886 82997: contig of 4012 bp in length
* 82998 83097: gap of unknown length
* 83098 86339: contig of 3242 bp in length
* 86340 89385: gap of unknown length
* 89386 89485: contig of 2946 bp in length
* 89486 93482: gap of unknown length
* 93483 96259: contig of 3897 bp in length
* 96260 99387: contig of 2777 bp in length
* 99388 100087: gap of unknown length
* 100088 103824: contig of 3628 bp in length
* 103825 109269: gap of unknown length
* 109270 112382: contig of 3737 bp in length
* 112383 112482: gap of unknown length
* 112483 116027: contig of 5345 bp in length
* 116028 121318: contig of 3545 bp in length
* 121319 121418: contig of 5191 bp in length
* 121419 125277: gap of unknown length
* 125278 125377: contig of 3859 bp in length
* 125378 129404: contig of 4027 bp in length
* 129405 129504: gap of unknown length
* 129505 135160: contig of 5656 bp in length
* 135161 135260: gap of unknown length
* 135261 141525: contig of 6265 bp in length
* 141526 141526: gap of unknown length
* 141527 147453: contig of 5828 bp in length

Query Match 14.4% Score 19; DB 2; Length 175037;
Best Local Similarity 100.08; Pred. No. 7.7;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 104 AGCAGGAGAACTGGAGTGA 122
Db 19685 AGCAGGAGAACTGGAGTGA 19703
|||||
***, 67 unordered pieces.

AC097955 176409 bp DNA linear HTG 11-JUL-2002
Rattus norvegicus clone CH230-110B17, *** SEQUENCING IN PROGRESS
***, 67 unordered pieces.

AC097955
AC097955 3 GI:21723867
HTG; HTGS_PHASE1.
Norway rat.
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 176409)
Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-Osman,F.R., Allen,C.,
Alsbrooks,S.L., Amarantunge,H.C., Are,J.R., Ayele,M., Banks,T.,
Barbaria,J., Benton,J., Bimage,K., Blankenburg,K., Bonnin,D.,
Bouck,J., Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P.,
Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,
Carron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,
Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,

```

Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,
 Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,
 Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,
 Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,
 Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,
 Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,
 Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,
 Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,
 Harris,C., Harris,K., Hart,M., Havlak,P., Haves,A., Hernandez,J.,
 Hernandez,O., Hodgson,A., Hognes,M., Holloway,C., Hollins,B.,
 Homsli,F., Howard,S., Huber,J., Hulyk,S., Hume,J., Jackson,L.E.,
 Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S.,
 Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C.,
 Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.,
 Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W., Lounseged,H.,
 Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,
 Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E.,
 Massey,E., Mawhney,E., McLeod,M.P., Meador,M., Mel,G., Metzker,M.,
 Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S.,
 Moser,M., Neal,D., Newton,J., Newton,N., Nguyen,A., Nguyen,S.,
 Nguyen,N., Nickerson,E., Nwokenkwo,S., Ogih,M., Okwuonu,G.,
 Ogunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L.,
 Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y.,
 Rives,M., Rojas,A., Rojebokan,I., Roife,M., Ruiz,S., Savary,G.,
 Scherer,S., Scott,G., Shen,H., Shooshtari,N., Sisson,I.,
 Sodergren,E., Sonaik,T., Sparks,A., Stanley,H., Stone,H.,
 Sutton,A., Svatek,A., Tabot,P., Tamerisa,A., Tamerisa,K., Tang,H.,
 Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S.,
 Usmani,K., Vasquez,L., Vera,V., Villalon,D., Vinson,R., Wang,Q.,
 Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S.,
 Williams,G., Williamson,A., Wleczyk,R., Wooden,S., Worley,K.,
 Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,
 Weinstock,G. and Gibbs,R.

Unpublished
 2 (bases 1 to 176409)
 Worley,K.C.
 Direct Submission
 Submitted (23-OCT-2001) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 3 (bases 1 to 176409)
 Worley,K.C.
 Direct Submission
 Submitted (11-JUL-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 On Jul 10, 2002 this sequence version replaced gi:17949462.

Center: Baylor College of Medicine
 Center code: BCM
 Web site: <http://www.hgsc.bcm.tmc.edu/>
 Contact: hgsc-help@bcm.tmc.edu
 Project Information
 Center project name: GFSE
 Center clone name: CH230-110B17
 Summary Statistics
 Sequencing vector: Plasmid;
 Chemistry: Dye-terminator Big Dye; 100% of reads
 Assembly program: Phrap; version 0.990329
 Consensus quality: 123062 bases at least Q40
 Consensus quality: 126575 bases at least Q30
 Consensus quality: 129664 bases at least Q20

 * NOTE: Estimated insert size may differ from sequence length
 * (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 67 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will

* be preserved.

```

1 1050: contig of 1050 bp in length
* 1051: gap of unknown length
* 1151: contig of 1062 bp in length
* 2212: gap of unknown length
* 2213: contig of 1222 bp in length
* 3534: gap of unknown length
* 3535: contig of 1065 bp in length
* 3635: gap of unknown length
* 4700: contig of 1070 bp in length
* 4800: gap of unknown length
* 5870: contig of 1208 bp in length
* 5970: gap of unknown length
* 7177: contig of 1216 bp in length
* 7278: gap of unknown length
* 8493: contig of 1206 bp in length
* 8593: gap of unknown length
* 8594: contig of 1206 bp in length
* 8999: gap of unknown length
* 9800: contig of 1448 bp in length
* 11347: gap of unknown length
* 11448: contig of 1249 bp in length
* 12697: gap of unknown length
* 12797: contig of 1596 bp in length
* 14393: gap of unknown length
* 14922: contig of 1188 bp in length
* 15880: gap of unknown length
* 15881: contig of 1167 bp in length
* 16948: gap of unknown length
* 17047: contig of 1173 bp in length
* 18221: gap of unknown length
* 18321: contig of 1001 bp in length
* 19421: gap of unknown length
* 19422: contig of 1161 bp in length
* 20582: gap of unknown length
* 20683: contig of 1569 bp in length
* 20683: gap of unknown length
* 22351: contig of 1018 bp in length
* 22352: gap of unknown length
* 23369: contig of 1659 bp in length
* 23469: gap of unknown length
* 25128: contig of 1659 bp in length
* 25129: gap of unknown length
* 25229: contig of 1971 bp in length
* 27200: gap of unknown length
* 27200: contig of 1507 bp in length
* 28806: gap of unknown length
* 28907: contig of 1667 bp in length
* 30573: gap of unknown length
* 30674: contig of 1398 bp in length
* 32072: gap of unknown length
* 32171: contig of 1668 bp in length
* 33839: gap of unknown length
* 33939: contig of 2399 bp in length
* 36338: gap of unknown length
* 36439: contig of 1575 bp in length
* 38013: gap of unknown length
* 38113: contig of 2267 bp in length
* 40380: gap of unknown length
* 40381: contig of 2539 bp in length
* 43019: gap of unknown length
* 43020: contig of 1523 bp in length
* 44642: gap of unknown length
* 44643: contig of 1857 bp in length
* 46599: gap of unknown length
* 46600: contig of 1921 bp in length
* 48620: gap of unknown length
* 48720: contig of 1628 bp in length
* 50348: gap of unknown length
* 50449: contig of 2259 bp in length
* 52707: gap of unknown length
* 52808: contig of 2308 bp in length
* 55116: gap of unknown length
* 55216: contig of 1533 bp in length
* 56748: gap of unknown length
* 56849: contig of 3107 bp in length
* 59555: gap of unknown length
* 60055: gap of unknown length

```

```

* 60056 61804: contig of 1749 bp in length
* 61805: gap of unknown length
* 61905: contig of 2173 bp in length
* 64078 64177: gap of unknown length
* 64178 67094: contig of 2917 bp in length
* 67095 67195: gap of unknown length
* 67195 68690: contig of 1496 bp in length
* 68691 68790: gap of unknown length
* 68791 71790: contig of 3000 bp in length
* 71791 71890: gap of unknown length
* 71891 75304: contig of 3414 bp in length
* 75305 75405: gap of unknown length
* 75405 78445: contig of 3041 bp in length
* 78446 78545: gap of unknown length
* 78546 81392: contig of 2847 bp in length
* 81393 81492: gap of unknown length
* 81493 83766: contig of 2274 bp in length
* 83767 83866: gap of unknown length
* 83867 86812: contig of 2946 bp in length
* 86813 86912: gap of unknown length
* 86913 88777: contig of 1965 bp in length
* 88778 88978: gap of unknown length
* 88979 91962: contig of 2985 bp in length
* 91963 92062: gap of unknown length
* 92063 95558: contig of 3496 bp in length
* 95559 95658: gap of unknown length
* 95659 99330: contig of 3672 bp in length
* 99331 99430: gap of unknown length
* 99431 103014: contig of 3584 bp in length
* 103015 103115: gap of unknown length
* 103115 106384: contig of 3270 bp in length
* 106385 106484: gap of unknown length
* 106485 109913: contig of 3429 bp in length
* 109914 110013: gap of unknown length
* 110014 113006: contig of 2993 bp in length

```

Query Match 14.4%; Score 19; DB 2; Length 176409;
 Best Local Similarity 100.0%; Pred. No. 7.7;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 104 ACCAGGAGAGTGGAGTGA 122
 Db 66235 AGCAGGAGAGTGGAGTGA 66217
 |||||

RESULT 44
 AC022742

LOCUS AC022742 Homo sapiens chromosome 15 clone RP11-762A1 map 15, WORKING DRAFT
 DEFINITION AC022742 Homo sapiens chromosome 15 clone RP11-762A1 map 15, WORKING DRAFT
 ACCESSION AC022742
 VERSION AC022742.4 GI:13123383
 KEYWORDS HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
 SOURCE Homo sapiens.

ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 176967)
 Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
 Anderson,S., Baldwin,J., Barna,N., Beckerly,R., Beda,F.,
 Boguslavskiy,L., Boukhgalter,B., Brown,A., Burkett,G., Castle,A.,
 Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cooke,P.,
 Dearellano,K., Dewar,K., Domino,M., Doyle,M., Fenesor,J.,
 Ferreira,P., FitzHugh,W., Forrest,C., Gage,D., Galagan,J.,
 Gardyna,S., Grant,G., Hagos,B., Heaford,A., Horton,L.,
 Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J.,
 Landers,T., Lechoczky,J., Levine,R., Lieu,C., Liu,G., Locke,K.,
 Macdonald,P., Marquis,N., McEwan,P., McGurk,A., McKernan,K.,
 McPheeters,R., Meldrim,J., Meneus,L., Morrow,J., Naylor,J.,

Norman, C.H., O'Connor, T., O'Donnell, P., Olivari, T.M., Peterson, K.,
 Pierre, N., Pisan, C., Pollara, V., Raymond, C., Riley, R., Rothman, D.,
 Roy, A., Santos, R., Severy, P., Spencer, B., Stange-Thomann, N.,
 Stojanovic, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J.,
 Tirrell, A., Vassiliev, H., Viel, R., Vo, A., Wu, X., Wyman, D., Ye, W.J.,
 Zimmer, A. and Zody, M.

TITLE JOURNAL

Submitted (06-FEB-2000) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 On Feb 25, 2001 this sequence version replaced gi:11990718.
 All repeats were identified using RepeatMasker:
 Smit, A.F.A. & Green, P. (1996-1997)
 http://ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR

Web site: http://www-seq.wi.mit.edu

Contact: sequence_submissions@genome.wi.mit.edu

----- Project Information

Center project name: L5677

Center clone name: 762_A1

----- Summary Statistics

Sequencing vector: M13; M7815; 2% of reads

Chemistry: Plasmid; n/a; 98% of reads

Assembly program: Phrap; version 0.960731

Consensus quality: 176032 bases at least Q40

Consensus quality: 176449 bases at least Q30

Consensus quality: 176615 bases at least Q20

Insert size: 186000; agarose-fp

Quality coverage: 9.0 in Q20 bases; agarose-fp

Quality coverage: 9.5 in Q20 bases.

* NOTE: This is a 'working draft' sequence. It currently
 * consists of 4 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

* 1 14400: contig of 14400 bp in length
 * 14401 14500: gap of 100 bp
 * 14501 61165: contig of 46665 bp in length
 * 61166 61265: gap of 100 bp
 * 61266 112521: contig of 51256 bp in length
 * 112522 112621: gap of 100 bp
 * 112622 176967: contig of 64346 bp in length.

FEATURES

source

1. 176967
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="15"
 /map="15"

/clone="RP11-762A1"
 /clone_lib="RPC1-11 Human Male BAC"

1. 14400

/note="assembly_fragment"

vector_end.SP6

14501..61165

/note="assembly_fragment"

61266..112521

/note="assembly_fragment"

112622..176967

/note="assembly_fragment"

vector_end.T7

vector_side:right"

300 others

BASE COUNT 53791 a 35746 c 36226 g 50904 t

ORIGIN

Query Match

Best Local Similarity

14.4%; Score 19; DB 2; Length 176967;

100.0%; Pred. No. 7.7;

Matches

19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

57 GCCCACAGAGATGCAGAG 75

|||||

Db 69605 GCCCACAGAGATGCAGAG 69623

RESULT 45

AC102783

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

AC102783 177033 bp DNA linear HTG 21-AUG-2002
 MUS musculus clone RP23-390K3, WORKING DRAFT SEQUENCE, 24 unordered
 pieces.
 AC102783.2 GI:22381794
 HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
 house mouse.
 Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 Birren, B., Nusbaum, C. and Lander, E.
 Mus musculus, clone RP23-390K3
 Unpublished
 2 (bases 1 to 177033)
 Birren, B., Linton, L., Nusbaum, C., Lander, E., Ali, A., Allen, N.,
 Anderson, S., Barna, N., Bastien, V., Boguslavsky, L., Boukhgalter, B.,
 Brown, A., Camarata, J., Campolano, A., Chang, J., Chazaro, B.,
 Choepel, Y., Colangelo, M., Collins, S., Collymore, A., Cook, A.,
 Cooke, P., DeArellano, K., Dewar, K., Diaz, J.S., Dodge, S., Faro, S.,
 Ferreira, P., FitzHugh, W., Gage, D., Galagan, J., Gardyna, S.,
 Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N.,
 Hagos, B., Heaford, A., Horton, L., Hulme, W., Iliev, I., Johnson, R.,
 Jones, C., Kamat, A., Karats, A., Kells, C., LaRocque, K.,
 Lamazares, R., Landers, T., Lehoczy, J., Levine, R., Liu, G.,
 MacLean, C., Macdonald, P., Major, J., Marquis, N., Matthews, C.,
 McCarthy, M., McEwan, P., McKernan, K., McPheeters, R., Meldrum, J.,
 Meneus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C.,
 Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neil, D.,
 Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V.,
 Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P.,
 Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schuback, R.,
 Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N.,
 Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J.,
 Topham, K., Travers, M., Travis, N., Triggillo, J., Vassiliev, H.,
 Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G.,
 Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.
 Direct Submission
 Submitted (23-NOV-2001) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 3 (bases 1 to 177033)
 Birren, B., Nusbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S.,
 Barna, N., Bastien, V., Bloom, T., Boguslavsky, L., Boukhgalter, B.,
 Camarata, J., Chang, J., Chazaro, B., Choepel, Y., Collymore, A.,
 Cook, A., Cooke, P., DeArellano, K., Dewar, K., Diaz, J.S., Dodge, S.,
 Faro, S., Ferreira, P., FitzGerald, M., Gage, D., Galagan, J.,
 Gardyna, S., Gord, S., Graham, L., Grand-Pierre, N., Hagos, B.,
 Horton, B., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A.,
 Karatas, A., Kells, C., Landers, T., Levine, R., Lindblad-Toh, K.,
 Liu, G., MacLean, C., Macdonald, P., Major, J., Matthews, C.,
 McCarthy, M., Meldrum, J., Meneus, L., Mihova, T., Mlenga, V.,
 Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C.H.,
 O'Connor, T., O'Donnell, P., O'Neil, D., Oliver, J., Peterson, K.,
 Phunkhang, P., Pierre, N., Raymond, C., Retta, R., Rise, C., Rogov, P.,
 Roman, J., Roy, A., Schauer, S., Schuback, R., Seaman, S., Severy, P.,
 Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Talamas, J.,
 Tesfaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H.,
 Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J.,
 Zembek, L., Zimmer, A. and Zody, M.
 Direct Submission
 Submitted (21-AUG-2002) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 On Aug 21, 2002 this sequence version replaced gi:17060893.
 All repeats were identified using RepeatMasker:

Smit, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>
 ----- Genome Center
 Center: Whitehead Institute/ MIT Center for Genome Research
 Center code: WIBR

Web site: <http://www-seq.wi.mit.edu>
 Contact: sequence_submissions@genome.wi.mit.edu
 ----- Project Information

Center project name: L19655
 Center clone name: 390_K_3
 ----- Summary Statistics

Sequencing vector: Plasmid; n/a; 100% of reads
 Chemistry: Dye-terminator Big Dye; 100% of reads
 Assembly program: Phrap; version 0.960731
 Consensus quality: 170563 bases at least Q40
 Consensus quality: 173449 bases at least Q30
 Consensus quality: 174139 bases at least Q20
 Insert size: 183000; agarose-fp
 Insert size: 174733; sum-of-contigs
 Quality coverage: 6.1 in Q20 bases; agarose-fp
 Quality coverage: 6.4 in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently
 * consists of 24 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

```

1
4612 4711: contig of 4611 bp in length
4712 5577: contig of 866 bp in length
5578 5677: gap of 100 bp
5678 6686: contig of 1009 bp in length
6687 6786: gap of 100 bp
6787 7249: contig of 463 bp in length
7250 7349: gap of 100 bp
7350 8617: contig of 1268 bp in length
8618 8717: gap of 100 bp
8718 10292: contig of 1575 bp in length
10293 10392: gap of 100 bp
10393 11672: contig of 1280 bp in length
11673 11772: gap of 100 bp
11773 13066: contig of 1294 bp in length
13067 13166: gap of 100 bp
13167 14607: contig of 1441 bp in length
14608 14707: gap of 100 bp
14708 17503: contig of 2796 bp in length
17504 17603: gap of 100 bp
17604 21653: contig of 4050 bp in length
21654 21753: gap of 100 bp
21754 25974: contig of 4221 bp in length
25975 26074: gap of 100 bp
26075 29880: contig of 3806 bp in length
29881 33369: contig of 3389 bp in length
33370 33469: gap of 100 bp
33470 70307: contig of 36838 bp in length
70308 70407: gap of 100 bp
70408 75527: contig of 5120 bp in length
75528 75627: gap of 100 bp
75628 84901: contig of 9274 bp in length
84902 85001: gap of 100 bp
85002 94921: contig of 9920 bp in length
94922 95021: gap of 100 bp
95022 100827: contig of 5806 bp in length
100828 100927: gap of 100 bp
100928 114503: contig of 13576 bp in length
114504 114603: gap of 100 bp
114604 135326: contig of 20723 bp in length
135327 135426: gap of 100 bp
135427 152074: contig of 16648 bp in length

```

```

* 152075 152174: gap of 100 bp
* 152175 174449: contig of 22275 bp in length
* 174450 174549: gap of 100 bp
* 174550 177033: contig of 2484 bp in length.

```

FEATURES

```

source
1..177033
  Location/Qualifiers
    /organism="Mus musculus"
    /db_xref="taxon:10090"
    /clone="RP23-390K3"
    /clone_lib="RPCI-23 Female Mouse BAC"
    1..4611
      /note="assembly_fragment"
      clone_end:SP6
      vector_side:left
      4712..5577
        /note="assembly_fragment"
      5678..6686
        /note="assembly_fragment"
      6787..7249
        /note="assembly_fragment"
      7350..8617
        /note="assembly_fragment"
      8718..10292
        /note="assembly_fragment"
      10393..11672
        /note="assembly_fragment"
      11773..13066
        /note="assembly_fragment"
      13167..14607
        /note="assembly_fragment"
      14708..17503
        /note="assembly_fragment"
      17604..21653
        /note="assembly_fragment"
      21754..25974
        /note="assembly_fragment"
      26075..29880
        /note="assembly_fragment"
      29981..33369
        /note="assembly_fragment"
      33470..70307
        /note="assembly_fragment"
      70408..75527
        /note="assembly_fragment"
      75628..84901
        /note="assembly_fragment"
      85002..94921
        /note="assembly_fragment"
      95022..100827
        /note="assembly_fragment"
      100928..114503
        /note="assembly_fragment"
      114604..135326
        /note="assembly_fragment"
      135427..152074
        /note="assembly_fragment"
      152175..174449
        /note="assembly_fragment"
      174550..177033
        /note="assembly_fragment"
      clone_end:T7
      vector_side:right

```

BASE COUNT 52386 a .33924 c 32870 g 55548 t 2305 others

Query Match 14.4%; Score 19; DB 2; Length 177033;
 Best Local Similarity 100.0%; Pred.No.7.7;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 49 AAGCTGAAGGCCACAGAGA 67

DB 138688 AAGCTGAAGGCCACAGAGA 138706

RESULT 46

AC018623
 LOCUS
 DEFINITION
 Homo sapiens chromosome 8 clone RP11-444C21 map 8, WORKING DRAFT
 SEQUENCE, 12 unordered pieces.
 AC018623
 AC018623.4 GI:7382432
 HTG: HTGS_PHASE1; HTGS_DRAFT.
 VERSION
 HTG: HTGS_PHASE1; HTGS_DRAFT.
 KEYWORDS
 Homo sapiens
 SOURCE
 Homo sapiens
 ORGANISM
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 180019)
 Birren,B., Linton,L., Nusbaum,C. and Lander,E.
 Homo sapiens chromosome 8, clone RP11-444C21
 Unpublished
 2 (bases 1 to 180019)
 Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
 Anderson,S., Baldwin,J., Barna,N., Beckerly,R., Bida,F.,
 Boguslavsky,L., Boukhgater,B., Brown,A., Burkett,G., Castle,A.,
 Choepel,Y., Colangelo,M., Collins,S., Collamore,A., Cooke,P.,
 Dearellano,K., Dewar,K., Domino,M., Doyle,M., Fenestor,J.,
 Ferreira,P., FitzHugh,W., Forrest,C., Gage,D., Galagan,J.,
 Gardina,S., Grant,G., Hagos,B., Heaford,A., Horton,L.,
 Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J.,
 Landers,T., Lehoczyk,J., Levine,R., Lieu,C., Liu,G., Locke,K.,
 MacDonald,P., Marquis,N., McEwan,P., McGurk,A., McKernan,K.,
 Meidlin,J., Meneus,L., Morrow,J., Naylor,J., Norman,C.H.,
 O'Connor,T., O'Donnell,P., Olivari,T.M., Peterson,K., Pierre,N.,
 Pisani,C., Pollara,V., Raymond,C., Riley,R., Rothman,D., Roy,A.,
 Santos,R., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
 Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J., Tirrell,A.,
 Vasiliev,H., Vlei,R., Vo,A., Wu,X., Wyman,D., Ye,W.J., Zimmer,A.
 and Zody,M.
 Direct Submission
 Submitted (14-DEC-1999) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 On Apr 1, 2000 this sequence version replaced gi:6653305.
 All repeats were identified using RepeatMasker:
 Smit, A.F.A. & Green, P. (1996-1997)
 http://ftp.genome.washington.edu/RM/RepeatMasker.html

 Center: Whitehead Institute/ MIT Center for Genome Research
 Center code: WIRB
 Web site: http://www-seq.wi.mit.edu
 Contact: sequence_submissions@genome.wi.mit.edu

 Project information
 Center project name: L5243
 Center clone name: 444_C_21

 Summary Statistics
 Sequencing vector: M13; M77815; 100% of reads
 Chemistry: Dye-terminator Big Dye; 100% of reads
 Assembly program: Phrap; version 0.960731
 Consensus quality: 168843 bases at least Q40
 Consensus quality: 173972 bases at least Q30
 Consensus quality: 176111 bases at least Q20
 Insert size: 180000; agarose-fp
 Insert size: 178919; sum-of-contigs
 Quality coverage: 4.8 in Q20 bases; agarose-fp
 Quality coverage: 4.8 in Q20 bases; sum-of-contigs

 NOTE: This is a 'working draft' sequence. It currently
 * consists of 12 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.
 *
 * 1 1358: contig of 1358 bp in length
 * 1359 1458: gap of 100 bp
 * 1459 3021: contig of 1563 bp in length
 * 3022 3121: gap of 100 bp

TITLE
JOURNAL

COMMENT

* 3122 7480: contig of 4359 bp in length
 * 7481 7580: gap of 100 bp
 * 7581 15678: contig of 12098 bp in length
 * 19679 19778: gap of 100 bp
 * 19779 27895: contig of 8117 bp in length
 * 27896 27995: gap of 100 bp
 * 27996 44261: contig of 16266 bp in length
 * 44262 44361: gap of 100 bp
 * 44362 60303: contig of 15942 bp in length
 * 60304 60403: gap of 100 bp
 * 60404 77029: contig of 16626 bp in length
 * 77030 77129: gap of 100 bp
 * 77130 95870: contig of 18741 bp in length
 * 95871 95970: gap of 100 bp
 * 95971 118459: contig of 22489 bp in length
 * 118460 118559: gap of 100 bp
 * 118560 145068: contig of 26509 bp in length
 * 145069 145168: gap of 100 bp
 * 145169 180019: contig of 34851 bp in length.
 Location/Qualifiers
 1. 180019
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /map="8"
 /chromosome="8"
 /clone="RP11-444C21"
 /clone_lib="RPC1-11 Human Male BAC"
 1. 1358
 /note="assembly_fragment"
 1459. 3021
 /note="assembly_fragment"
 /note="assembly_fragment"
 clone_end:SP6
 vector_side:right
 3122. 7480
 /note="assembly_fragment"
 7581. 19678
 /note="assembly_fragment"
 19779. 27895
 /note="assembly_fragment"
 clone_end:T7
 27996. 44261
 /note="assembly_fragment"
 44362. 60303
 /note="assembly_fragment"
 60404. 77029
 /note="assembly_fragment"
 77130. 95870
 /note="assembly_fragment"
 95971. 118459
 /note="assembly_fragment"
 118560. 145068
 /note="assembly_fragment"
 145169. 180019
 /note="assembly_fragment"
 a 30783 c 32284 g 57225 t 1101 others

FEATURES

source

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

RESULT 47

AC110005
 LOCUS
 DEFINITION
 Homo sapiens chromosome 5 clone RP11-265K23, complete sequence.
 ACCESSION
 AC110005
 VERSION
 AC110005.2 GI:20279403

Query Match 14.4%; Score 19; DB 2; Length 180019;
 Best Local Similarity 100.0%; Pred No. 7.7;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 112 AAGTGGAGTGAATTCCT 130

Db 12396 AAGTGGAGTGAATTCCT 12414

* 25667 34242: contig of 8576 bp in length
 * 34243 34342: gap of 100 bp
 * 34343 98428: contig of 64087 bp in length
 * 98430 98529: gap of 100 bp
 * 98530 117192: contig of 18663 bp in length
 * 117193 117292: gap of 100 bp
 * 117293 133546: contig of 16254 bp in length
 * 133547 133646: gap of 100 bp
 * 133647 177342: contig of 43696 bp in length
 * 177343 177442: gap of 100 bp
 * 177443 185027: contig of 7585 bp in length.

FEATURES

source
 1. 185027 Location/Qualifiers
 /organism="Mus musculus"
 /db_xref="taxon:10090"
 /clone="RP24-396N7"
 /clone_lib="RPCi-24 Male Mouse BAC"
 misc_feature
 1. 2482
 /note="assembly_fragment"
 clone_end:SP6
 vector_side:left
 2583..13026
 /note="assembly_fragment"
 13127..25566
 /note="assembly_fragment"
 25667..34242
 /note="assembly_fragment"
 34343..98429
 /note="assembly_fragment"
 98530..117192
 /note="assembly_fragment"
 117293..133546
 /note="assembly_fragment"
 133647..177342
 /note="assembly_fragment"
 177443..185027
 /note="assembly_fragment"
 clone_end:T7
 vector_side:right
 59027 a 33873 c 33903 g 57424 t 800 others

Query Match 14.4%; Score 19; DB 2; Length 185027;
 Best Local Similarity 100.0%; Pred. No. 7.7;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 63 AGAGATGCAGAGAACACT 81

Db 122180 AGAGATGCAGAGAACACT 122198

RESULT 49
 AC116058/c
 LOCUS
 DEFINITION Rattus norvegicus clone CH230-204F1, *** SEQUENCING IN PROGRESS
 *** 70 unordered pieces.

AC116058
 AC116058.2 Gi:21732014

VERSION HTG; HTGS_PHASE1.

KEYWORDS Norway rat.

SOURCE Rattus norvegicus

ORGANISM Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

REFERENCE 1 (bases 1 to 185828)

AUTHORS Muzny,D.M., Adams,C., Adlo-Oduola,B., Ali-osman,F.R., Allen,C., Alsbrooks,S.L., Amarantunge,H.C., Are,J.R., Ayele,M., Banks,T., Barbara,J., Benton,J., Bimage,K., Blankenburg,K., Bonnin,D., Bouck,J., Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P., Buhay,C., Burch,P., Burkett,C., Burrell,K.B., Byrd,N.C., Carron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C., Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,

Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A., Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H., Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J., Earhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M., Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P., Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R., Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K., Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J., Hernandez,O., Hodgson,A., Hogues,M., Holloway,C., Hollins,B., Homsí,F., Howard,S., Huber,J., Hulyk,S., Hume,J., Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S., Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C., Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L., Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W., Louiseged,H., Lozadó,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J., Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E., Massey,E., Mawhiney,E., McLeod,M.P., Meador,M., Mei,G., Metzker,M., Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S., Moser,M., Neal,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N., Nguyen,N., Nickerson,E., Nwokenkwo,S., Oguh,M., Okwuonu,G., Ozagunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L., Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y., Rives,M., Rojas,A., Rojibokan,I., Rolfe,M., Ruiz,S., Savary,G., Scherer,S., Scott,G., Shen,H., Shoohtari,N., Sisson,I., Sodergren,E., Sonaik,T., Sparks,A., Stanley,H., Stone,H., Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H., Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S., Usmani,K., Vasquez,L., Vera,V., Villalón,D., Vinson,R., Wang,Q., Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S., Williams,G., Williamson,A., Wleczyk,R., Wooden,S., Worley,K., Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D., Weinstock,G. and Gibbs,R.

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Center: Baylor College of Medicine

Center code: BCM

Web site: <http://www.hgsc.bcm.tmc.edu/>

Contact: hgsc-help@bcm.tmc.edu

----- Project Information

Center project name: GLKJ

Center clone name: CH230-204F1

----- Summary Statistics

Sequencing vector: Plasmid;

Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.990329

Consensus quality: 127424 bases at least Q40

Consensus quality: 131027 bases at least Q30

Consensus quality: 133711 bases at least Q20

NOTE: Estimated insert size may differ from sequence length
 (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
 NOTE: This is a 'working draft' sequence. It currently consists of 70 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

* 1433: contig of 1433 bp in length
* 1434: 1533: gap of unknown length
* 1534: 2923: contig of 1390 bp in length
* 2924: 3023: gap of unknown length
* 3024: 4464: contig of 1441 bp in length
* 4465: 4564: gap of unknown length
* 4565: 5999: contig of 1435 bp in length
* 6000: 6999: gap of unknown length
* 6999: 7803: contig of 1704 bp in length
* 7804: 7903: gap of unknown length
* 7904: 9065: contig of 1162 bp in length
* 9066: 9165: gap of unknown length
* 9166: 10363: contig of 1198 bp in length
* 10364: 10463: gap of unknown length
* 10464: 11640: contig of 1177 bp in length
* 11641: 11740: gap of unknown length
* 11741: 12761: contig of 1021 bp in length
* 12762: 12861: gap of unknown length
* 12862: 14266: contig of 1405 bp in length
* 14267: 14366: gap of unknown length
* 14367: 15835: contig of 1469 bp in length
* 15836: 15935: gap of unknown length
* 15936: 16949: contig of 1014 bp in length
* 16950: 17049: gap of unknown length
* 17050: 18237: contig of 1188 bp in length
* 18238: 18337: gap of unknown length
* 18338: 19787: contig of 1450 bp in length
* 19788: 19887: gap of unknown length
* 19888: 21086: contig of 1199 bp in length
* 21087: 21186: gap of unknown length
* 21187: 22879: contig of 1693 bp in length
* 22880: 22979: gap of unknown length
* 22980: 24058: contig of 1079 bp in length
* 24059: 24158: gap of unknown length
* 24159: 25441: contig of 1283 bp in length
* 25442: 25541: gap of unknown length
* 25542: 26933: contig of 1392 bp in length
* 26934: 27033: gap of unknown length
* 27034: 28060: contig of 1027 bp in length
* 28061: 28160: gap of unknown length
* 28161: 29434: contig of 1274 bp in length
* 29435: 29534: gap of unknown length
* 29535: 30838: contig of 1304 bp in length
* 30839: 30938: gap of unknown length
* 30939: 32455: contig of 1517 bp in length
* 32456: 32555: gap of unknown length
* 32556: 34004: contig of 1449 bp in length
* 34005: 34104: gap of unknown length
* 34105: 35351: contig of 1247 bp in length
* 35352: 35451: gap of unknown length
* 35452: 36813: contig of 1362 bp in length
* 36814: 36913: gap of unknown length
* 36914: 37961: contig of 1048 bp in length
* 37962: 38061: gap of unknown length
* 38062: 40261: contig of 2200 bp in length
* 40262: 40361: gap of unknown length
* 40362: 42319: contig of 1858 bp in length
* 42320: 42319: gap of unknown length
* 42320: 44732: contig of 2413 bp in length
* 44733: 44832: gap of unknown length
* 44833: 45994: contig of 1162 bp in length
* 45995: 46094: gap of unknown length
* 46095: 47638: contig of 1544 bp in length
* 47639: 47738: gap of unknown length
* 47739: 50033: contig of 2295 bp in length
* 50034: 50133: gap of unknown length
* 50134: 51982: contig of 1849 bp in length
* 51983: 52082: gap of unknown length
* 52083: 54517: contig of 2435 bp in length
* 54518: 54617: gap of unknown length
* 54618: 55793: contig of 1176 bp in length
* 55794: 55893: gap of unknown length
* 55894: 57369: contig of 1476 bp in length

* 57370 57469: gap of unknown length
* 57470 60281: contig of 2812 bp in length
* 60282 60381: gap of unknown length
* 60382 63015: contig of 2634 bp in length
* 63016 63115: gap of unknown length
* 63116 64232: contig of 1117 bp in length
* 64233 64332: gap of unknown length
* 64333 66937: contig of 2505 bp in length
* 66938 66937: gap of unknown length
* 66938 68774: contig of 1837 bp in length
* 68775 68874: gap of unknown length
* 68875 70650: contig of 1776 bp in length
* 70651 70750: gap of unknown length
* 70751 72860: contig of 2110 bp in length
* 72861 72961: gap of unknown length
* 72961 74772: contig of 1812 bp in length
* 74773 74872: gap of unknown length
* 74873 77156: contig of 2284 bp in length
* 77157 77256: gap of unknown length
* 77257 79711: contig of 2455 bp in length
* 79712 79811: gap of unknown length
* 79812 82322: contig of 2511 bp in length
* 82323 82422: gap of unknown length
* 82423 85227: contig of 2805 bp in length
* 85228 85327: gap of unknown length
* 85328 87271: contig of 1944 bp in length
* 87272 87371: gap of unknown length
* 87372 90074: contig of 2703 bp in length
* 90075 90174: gap of unknown length
* 90175 93369: contig of 3195 bp in length
* 93370 93469: gap of unknown length
* 93470 96888: contig of 3419 bp in length
* 96889 96988: gap of unknown length
* 96989 100692: contig of 3704 bp in length

Query Match 14.4%; Score 19; DB 2; Length 185828;
Best Local Similarity 100.0%; Pred. No. 7.7;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 97 ACCACAGACGAGGAGT 115
Db 9198 ACCACAGACGAGGAGT 9180

RESULT 50

AC020892

LOCUS AC020892 186253 bp DNA linear PRI 04-APR-2001
DEFINITION Homo sapiens chromosome 15 clone RP11-313P18 map 15q21.2, complete sequence.

ACCESSION AC020892

VERSION AC020892.7 GI:13540717

KEYWORDS HTG.

SOURCE Homo sapiens.

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

1 (bases 1 to 186253)

Rowen, L., Madan, A., Qin, S., Baradarani, L., Birditt, B., Bloom, S.,

Burke, J., Dors, M., Fleetwood, P., Kaur, A., Madan, A., Nesbitt, R.,

Pate, D., and Hood, L.

Sequencing of human chromosome 15 D15S146-D15S117 region

Unpublished

2 (bases 1 to 186253)

Rowen, L., Madan, A., Qin, S., Abbasi, N., Baradarani, L., Birditt, B.,

Bloom, S., Dors, M., Dickhoff, R., Fleetwood, P., Harrison, G.,

James, R., Kaur, A., Madan, A., Owen, M. P., Ratcliffe, A., Shaffer, T.

and Hood, L.

Direct Submission

Submitted (12-JAN-2000) Multimegabase Sequencing Center, University

of Washington, PO BOX 357730, Seattle, WA 98195, USA

3 (bases 1 to 186253)

Rowen, L., Madan, A., Qin, S., Baradarani, L., Birditt, B., Bloom, S.,

Burke, J., Dors, M., Fleetwood, P., Kaur, A., Madan, A., Nesbitt, R.,

TITLE Pate, D. and Hood, L.
 JOURNAL Direct Submission
 Submitted (04-APR-2001) Multimegabase Sequencing Center, Institute
 for Systems Biology, 4225 Roosevelt Way NE, Suite 200, Seattle, WA
 98105, USA
 COMMENT On Apr. 4, 2001 this sequence version replaced gi:13162486.

----- Genome Center
 Center: Multimegabase Sequencing Center
 Center code: UWMSC
 Web site: http://chroma.mbt.washington.edu/msg_www
 Contact: leerowen@systemsbiology.org
 ----- Summary Statistics
 Sequencing vector: pUC18; L08752
 Chemistry: Dye-terminator Big Dye; 90% of reads
 Chemistry: Dye-primer Big Dye; 10% of reads
 Assembly program: Phrap; version 0.990399

 Note: Data from overlapping BACs AC066613 [drafting center UWMSC],
 AC020685 [drafting center WIBR], and AC026770 [drafting center
 UWMSC] was added for finishing.

FEATURES
 Source 1..186253
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="15"
 /map="15q21.2"
 /clone="RP11-313P18"
 /clone.lib="RPC1 human BAC library 11"
 /note="Note: Data from overlapping BACs #RP11-707P17,
 RP11-414J1, and CTD-2308G16 was added and the consensus
 sequence was determined from RP11-313P18 to the extent
 possible"
 misc_feature 1..2529
 /note="Overlap with RP11-707P17 AC066613"
 unsure 9006..9025
 /note="low quality data"
 unsure 9945..9950
 /note="low quality data"
 unsure 14447
 /note="low quality data"
 unsure 14690..14695
 /note="low quality data"
 unsure 23515..23525
 /note="low quality data"
 unsure 24918..24925
 /note="low quality data"
 unsure 49155..49160
 /note="low quality data"
 unsure 90499..90505
 /note="low quality data"
 unsure 91255..91350
 /note="low quality data"
 misc_feature 93480..186253
 /note="Overlap with clone RP11-414J1 AC020685"
 misc_feature 176791..186253
 /note="Overlap with clone CTD-2308G16 AC026770"
 BASE COUNT 60025 a 36875 c 35294 g 54059 t
 ORIGIN

Query Match 14.48; Score 19; DB 9; Length 186253;
 Best local Similarity 100.0%; Pred. No. 7.7;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 57 GGCCACAGATGCAGAG 75
 Db 163450 GGCCACAGATGCAGAG 163468
 |||||||||||||||

Search completed: June 2, 2003, 16:15:05
 Job time : 1159 secs

GenCore version 5.1.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 2, 2003, 15:39:40 ; Search time 1428 Seconds
(without alignments)
1497.062 Million cell updates/sec

Title: US-09-915-178-1

Perfect score: 132

Sequence: 1 atggcacacaaactagacct.....agtggagtgaatttctctga 132

Scoring table: OLIGO_NUC

Gapop 60.0 , Gapext 60.0

Searched: 16154066 seqs, 8097743376 residues

Word size : 10

Total number of hits satisfying chosen parameters: 2551178

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database :

EST:*

1: em_estba:*

2: em_esthum:*

3: em_estin:*

4: em_estnu:*

5: em_estov:*

6: em_estpl:*

7: em_estro:*

8: em_htc:*

9: gb_est1:*

10: gb_est2:*

11: gb_htc:*

12: gb_est3:*

13: gb_est4:*

14: gb_est5:*

15: em_estfun:*

16: em_estom:*

17: gb_gss:*

18: em_gss_hum:*

19: em_gss_inv:*

20: em_gss_pln:*

21: em_gss_vrt:*

22: em_gss_fun:*

23: em_gss_nam:*

24: em_gss_mus:*

25: em_gss_other:*

26: em_gss_pro:*

27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	25	18.9	374	10	AW620090 852 MARC
2	21	15.9	382	12	BG410966 EM1_28_B0
3	21	15.9	418	13	BM125588
4	20	15.2	608	17	AZ098162 RPI-23-1
5	20	15.2	732	17	AG012646 Homo sapi
6	20	15.2	997	13	B1524256 603052058

19	14.4	198	14	N88692	N88692 K5134F Huma
19	14.4	203	12	BF800160	BF800160 PMO-C1004
19	14.4	354	17	A2808724	A2808724 2M0072624
19	14.4	509	17	A2250549	A2250549 RPI-23-1
19	14.4	530	13	B0084935	B0084935 B0084935
19	14.4	629	17	A2769784	A2769784 1M0570M13
19	14.4	655	14	BQ803203	BQ803203 WHE2834_H
19	14.4	678	12	BF317304	BF317304 601904069
19	14.4	724	12	BE882559	BE882559 601507007
19	14.4	1058	17	CNS052GR	AL318132 Tetraodon
19	14.4	135	9	AA096439	AA096439 19794.seq
19	13.6	144	9	AA623585	AA623585 vn73d02.r
19	13.6	154	9	AI156792	AI156792 ue54f03.r
19	13.6	157	9	AA967821	AA967821 uh05f03.r
19	13.6	157	14	D76516	D76516 MUS70C08 mo
19	13.6	158	9	AA792584	AA792584 vs88e04.r
19	13.6	158	14	W91170	W91170 mf71f03.r1
19	13.6	160	9	AA270923	AA270923 va70f05.r
19	13.6	161	9	AA207933	AA207933 mv92c01.r
19	13.6	162	9	AA792947	AA792947 vp28e01.r
19	13.6	162	9	AA220778	AA220778 mv69c04.r
19	13.6	162	9	AA591963	AA591963 v150b08.r
19	13.6	163	14	BM843852	BM843852 K-EST0121
19	13.6	164	14	BM686528	BM686528 UI-E-C00-
19	13.6	169	9	AA144770	AA144770 mr68a05.r
19	13.6	169	12	BE774336	BE774336 MRI-UM000
19	13.6	170	14	R57257	R57257 F2310 Fetal
19	13.6	171	9	AA375556	AA375556 EST87853
19	13.6	172	14	W18665	W18665 mb98604.r1
19	13.6	173	9	AA692385	AA692385 vt21a10.r
19	13.6	174	14	T27818	T27818 EST17484 Hu
19	13.6	179	12	BE774172	BE774172 MRI-UM000
19	13.6	180	9	AA068081	AA068081 mm56e02.r
19	13.6	180	10	AW647096	AW647096 EST325639
19	13.6	181	9	AA566472	AA566472 vm48b04.r
19	13.6	182	9	AA316715	AA316715 EST188545
19	13.6	184	14	N83763	N83763 KK2711F Hum
19	13.6	187	10	AW211038	AW211038 uc90d11.y
19	13.6	189	9	AA072898	AA072898 mm72d12.r
19	13.6	195	9	AA087727	AA087727 mol3e09.r
19	13.6	195	9	AA454807	AA454807 zx77f10.r
19	13.6	198	9	AV044717	AV044717 AV044717
19	13.6	198	10	BE168674	BE168674 QV1-HT051
19	13.6	201	9	AA302441	AA302441 EST15672
19	13.6	201	14	W16355	W16355 mb556c07.r1
19	13.6	203	9	AA403933	AA403933 va32a01.r
19	13.6	203	9	AA467283	AA467283 vf06e09.r
19	13.6	203	9	AA538208	AA538208 vj56h09.r
19	13.6	205	14	H92399	H92399 ys86a06.r1
19	13.6	207	9	AA312784	AA312784 EST183467
19	13.6	209	9	AA416035	AA416035 vf40h12.r
19	13.6	210	9	AA301761	AA301761 EST14990
19	13.6	210	9	AA623434	AA623434 vn33a11.r
19	13.6	212	9	AA445744	AA445744 vf09e09.r
19	13.6	212	9	AA472510	AA472510 vho7d12.r
19	13.6	213	9	AA140326	AA140326 mr66e10.r
19	13.6	213	9	AA207455	AA207455 mv87c04.r
19	13.6	213	14	F26549	F26549 HSPD14071 H
19	13.6	214	14	W89298	W89298 mf63c09.r1
19	13.6	215	9	AA666996	AA666996 vk87h04.r
19	13.6	215	9	AA546784	AA546784 vq02d02.r
19	13.6	215	10	AW369977	AW369977 ILO-BT023
19	13.6	216	14	T92656	T92656 ye22e05.r1
19	13.6	216	17	AQ545403	AQ545403 CITBI-E1-
19	13.6	217	12	BF041410	BF041410 BP250008B
19	13.6	217	12	BF083159	BF083159 QVO-CS001
19	13.6	217	14	D53771	D53771 HUM119F07B
19	13.6	218	9	AA137562	AA137562 mq37c10.r
19	13.6	221	10	AW886615	AW886615 RCI-OT008
19	13.6	223	14	N86757	N86757 L0512F Huma
19	13.6	226	10	AW800960	AW800960 MR3-UM006
19	13.6	227	14	T60895	T60895 yb72f05.r1
19	13.6	229	9	AA091048	AA091048 cchn1095.

80 18 13.6 230 9 AA382739
81 18 13.6 230 14 C06531
82 18 13.6 232 9 AA095816
83 18 13.6 234 12 BF924141
84 18 13.6 237 9 AA372460
85 18 13.6 239 9 AA340512
86 18 13.6 239 9 AA366141
87 18 13.6 239 10 AW800966
88 18 13.6 240 9 AA302239
89 18 13.6 240 14 T25985
90 18 13.6 241 9 AA302528
91 18 13.6 241 10 AW998269
92 18 13.6 241 10 AW998418
93 18 13.6 242 10 AW998418
94 18 13.6 243 9 AU059537
95 18 13.6 243 10 AW797675
96 18 13.6 243 10 AW797688
97 18 13.6 244 12 BF012761
98 18 13.6 244 14 BQ840431
99 18 13.6 245 14 W14700
100 18 13.6 247 9 AA358097
101 18 13.6 247 9 AA362157
102 18 13.6 247 13 BG997026
103 18 13.6 251 9 AA301772
104 18 13.6 252 9 AA374250
105 18 13.6 252 14 H40563
106 18 13.6 253 10 AW797680
107 18 13.6 254 14 N85477
108 18 13.6 255 9 AL118094
109 18 13.6 255 9 AA617515
110 18 13.6 257 9 AA354418
111 18 13.6 257 10 AW800885
112 18 13.6 258 9 AA307117
113 18 13.6 258 9 AA339727
114 18 13.6 258 10 AW794138
115 18 13.6 260 13 BG994832
116 18 13.6 260 9 AA365696
117 18 13.6 261 10 AW998455
118 18 13.6 261 10 AW797742
119 18 13.6 263 9 AA316198
120 18 13.6 264 12 BG814678
121 18 13.6 266 10 AW801054
122 18 13.6 266 14 W11112
123 18 13.6 267 12 BE714634
124 18 13.6 268 9 AA685446
125 18 13.6 270 10 AW998366
126 18 13.6 270 9 AA304229
127 18 13.6 271 9 AA354636
128 18 13.6 271 14 BQ309112
129 18 13.6 272 14 W01438
130 18 13.6 273 9 AA301890
131 18 13.6 273 10 AW797766
132 18 13.6 273 10 AW800972
133 18 13.6 274 9 AA308404
134 18 13.6 274 10 AW797701
135 18 13.6 274 13 B1496663
136 18 13.6 274 14 D54136
137 18 13.6 276 9 AA365522
138 18 13.6 276 14 BM855353
139 18 13.6 279 9 AA352157
140 18 13.6 279 9 AA362030
141 18 13.6 280 9 AA022711
142 18 13.6 280 10 AW797700
143 18 13.6 282 9 AA301934
144 18 13.6 285 12 BE714668
145 18 13.6 287 10 AW793109
146 18 13.6 288 9 AA338821
147 18 13.6 289 9 AA374209
148 18 13.6 289 9 AA616956
149 18 13.6 290 9 AA070672
150 18 13.6 290 9 A1265376
151 18 13.6 290 9 AA380802
152 18 13.6 291 10 BB564596
153 18 13.6 292 9 AA304555
154 18 13.6 292 14 D56366
155 18 13.6 293 9 AA301999
156 18 13.6 293 10 AW619458
157 18 13.6 295 9 AA384610
158 18 13.6 296 13 BM125864
159 18 13.6 296 14 D53023
160 18 13.6 297 9 AA073962
161 18 13.6 297 9 AV136898
162 18 13.6 299 9 AA340452
163 18 13.6 299 14 BM855208
164 18 13.6 300 9 AU098753
165 18 13.6 300 9 AA316498
166 18 13.6 300 9 AA373192
167 18 13.6 300 10 AW325404
168 18 13.6 301 14 R11297
169 18 13.6 302 9 A1157676
170 18 13.6 302 9 AA309673
171 18 13.6 302 10 AW800956
172 18 13.6 303 9 AA316425
173 18 13.6 303 9 AA615691
174 18 13.6 303 14 D55279
175 18 13.6 304 9 AA657158
176 18 13.6 304 9 AA686446
177 18 13.6 304 9 AA745537
178 18 13.6 304 14 F29410
179 18 13.6 305 9 AA840447
180 18 13.6 306 12 BF884106
181 18 13.6 307 9 AA371161
182 18 13.6 309 9 AV120605
183 18 13.6 309 9 AA360865
184 18 13.6 309 9 AA372704
185 18 13.6 310 9 AA661462
186 18 13.6 310 9 AA303888
187 18 13.6 310 14 BQ367731
188 18 13.6 311 10 AW890346
189 18 13.6 312 14 F32364
190 18 13.6 312 14 F36325
191 18 13.6 312 14 H23750
192 18 13.6 313 9 AL034856
193 18 13.6 313 9 AA304699
194 18 13.6 313 14 W19791
195 18 13.6 314 9 AA138347
196 18 13.6 314 9 AA304702
197 18 13.6 314 14 C17724
198 18 13.6 314 14 D55359
199 18 13.6 317 10 AW801124
200 18 13.6 317 9 AA340143
201 18 13.6 318 12 BE863715
202 18 13.6 318 14 BM855196
203 18 13.6 319 9 AA304443
204 18 13.6 319 10 AW619459
205 18 13.6 321 9 AA092043
206 18 13.6 322 9 AA686035
207 18 13.6 322 14 BQ400156
208 18 13.6 322 14 D54437
209 18 13.6 322 14 F24677
210 18 13.6 323 9 AA870213
211 18 13.6 324 9 AA352586
212 18 13.6 324 9 AA590964
213 18 13.6 325 14 W25668
214 18 13.6 326 9 AA307549
215 18 13.6 326 14 W82636
216 18 13.6 327 10 AW945234
217 18 13.6 329 9 A1048068
218 18 13.6 329 14 H41377
219 18 13.6 330 9 AA315972
220 18 13.6 330 9 AA316512
221 18 13.6 330 10 AW869624
222 18 13.6 331 9 AL363867
223 18 13.6 331 9 AL363867
224 18 13.6 333 14 BQ324259
225 18 13.6 335 9 AA062353
AA382739 EST96282
C06531 C06531 Rat
AA095816 16004.seq
BF924141 MR4-NT014
AA372460 EST84358
AA340512 EST45729
AA366141 EST77049
AW800966 MR3-UM006
AA302239 EST15348
T25985 ATH467 HTCD
AA302528 EST15689
AW998269 PM2-BN006
AW998418 PM2-BN006
AU059537 AU059537
AW797675 CM0-UM004
AW797688 CM0-UM004
BF012761 ux73c01.y
BQ840431 mah67g10.
W14700 mb34c04.r1
AA358097 EST66956
AA362157 EST71636
BG997026 MR4-HT105
AA301772 EST14827
AA374250 EST86362
H40563 yn87c12.r1
AW797680 CM0-UM004
N85477 J3657F Huma
AL118094 r9225a26
AA617515 v177g01.r
AA354418 EST62702
AW800885 MR3-UM006
AA307117 EST17803
AA339727 EST44841
AW794138 MR1-UM000
BG994832 PM0-HT116
AA365696 EST76526
AW998455 PM2-BN006
AW797742 CM2-UM004
AA316198 EST187913
BG814678 dab05f10.
BF378584 CM0-UM004
AW801054 MR3-UM006
W11112 ma78h09.r1
BE714634 PM2-HT072
AA685446 EST107195
AW998366 PM2-BN006
AA304229 EST16960
AA354636 EST62927
BQ309112 MR0-BT600
W01438 za73c04.r1
AA301890 EST14940
AW797766 CM2-UM004
AW800972 MR3-UM006
AA308404 EST179234
AW797701 CM0-UM004
B1496663 df127b03.
D54136 HDM129G11B
AA365522 EST76310
BM855353 K-EST0138
AA352157 EST60371
AA362030 EST1374
AA022711 ze71f08.r
AW797700 CM0-UM004
AA301934 EST15013
BE714668 PM2-HT072
AW793109 CM4-UM000
AA338821 EST43818
AA374209 EST86338
AA616956 v120f03.r
AA070672 2m3g03.r
A1265376 uk03h11.y
BB564596 BB564596

c 226 18 13.6 337 10 AW415710
227 18 13.6 339 9 AL117845
228 18 13.6 339 14 D55927
229 18 13.6 340 9 AA108087
230 18 13.6 340 9 AA305923
231 18 13.6 340 9 AA316779
232 18 13.6 340 12 BF583502
233 18 13.6 342 14 W11627
c 234 18 13.6 343 9 AA945733
235 18 13.6 344 9 AA352601
236 18 13.6 344 14 F28066
237 18 13.6 344 14 H35100
c 238 18 13.6 345 9 AI952139
c 239 18 13.6 348 14 D51057
240 18 13.6 349 9 AA317870
241 18 13.6 349 10 AW824021
242 18 13.6 350 9 AA089668
243 18 13.6 350 14 F27686
244 18 13.6 350 14 W83798
245 18 13.6 352 14 T62988
246 18 13.6 353 9 AA709866
247 18 13.6 353 10 AW619828
248 18 13.6 354 9 AA313946
249 18 13.6 354 14 W62142
250 18 13.6 355 9 AA821726
251 18 13.6 355 14 D52330
252 18 13.6 357 12 BF934709
253 18 13.6 357 14 T51439
254 18 13.6 358 9 AA675439
c 255 18 13.6 358 9 AA843392
256 18 13.6 358 9 AA486145
257 18 13.6 358 14 T49877
c 258 18 13.6 359 9 AA821760
259 18 13.6 359 9 AI626080
260 18 13.6 359 10 BE178823
c 261 18 13.6 360 9 AI718100
262 18 13.6 360 9 AO097775
c 263 18 13.6 360 9 AA407170
c 264 18 13.6 361 9 AI122585
c 265 18 13.6 361 10 AW415709
c 266 18 13.6 361 14 T81560
c 267 18 13.6 362 9 AI468234
268 18 13.6 362 14 H56367
269 18 13.6 362 12 BE833110
270 18 13.6 364 9 AA869404
271 18 13.6 364 9 AA410037
272 18 13.6 364 14 BM855238
c 273 18 13.6 365 9 AA051064
c 274 18 13.6 366 13 B4499546
275 18 13.6 366 14 T49863
276 18 13.6 366 14 T59199
c 277 18 13.6 368 9 AA746615
278 18 13.6 368 12 BF687930
279 18 13.6 368 14 W10194
280 18 13.6 369 13 BG982434
c 281 18 13.6 369 14 F28020
c 282 18 13.6 370 9 AI707460
283 18 13.6 370 14 BM768418
c 284 18 13.6 370 14 BM768693
c 285 18 13.6 371 9 AI024403
c 286 18 13.6 371 9 AA407038
c 287 18 13.6 371 9 AA523332
c 288 18 13.6 372 9 AI761042
c 289 18 13.6 372 9 AA410036
c 290 18 13.6 372 9 AA574047
c 291 18 13.6 373 9 AA644882
c 292 18 13.6 373 10 BE620722
c 293 18 13.6 373 12 BE833193
c 294 18 13.6 374 9 AI718994
c 295 18 13.6 375 9 AA500732
c 296 18 13.6 375 10 AW262820
c 297 18 13.6 376 9 AI153298
c 298 18 13.6 376 9 AI708841
AW415710 50162 MAR
AL117845 m9112a16
D55927 HUM406E07B
AA108087 m158d08.r
AA305923 EST17707
AA316779 EST188447
BF583502 602101532
W11627 mbl7a01.r1
AA945733 EST201232
AA352601 EST60618
F28066 HSPDL6523 H
H35100 EST105990 R
AI952139 wx48e12.x
D51057 HUM009F08A
AA317870 EST19886
AW824021 uf62f02.y
AA089668 chn0395.s
F27686 HSPDL5764 H
W83798 mf33c06.r1
T62988 yb9a01.r1
AA709866 vt35a07.r
AW619828 7898 MARC
AA313946 EST185803
W62142 md86g04.r1
AA821726 vp71d10.s
D52330 HUM074H02B
BF934709 IL2-NT020
T51439 yb24a04.r1
AA675439 vr01d07.r
AA843392 aj16g11.s
AA486145 ab14c11.r
T49877 ya87g07.r1
AA821760 vp77c11.r
AI626080 ar87g06.x
BE178823 PM0-HR060
AI718100 as49b02.x
AO097775 AO097775
AA407170 EST02251
AI122585 qa88c06.x
AW415709 50161 MAR
T81560 yd28c11.r1
AI468234 tg34h09.x
H56367 yq97g10.r1
BE833110 QV3-OT006
AA869404 vq07f07.r
AA410037 EST01913
BM855238 K-EST0138
AA051064 mg73c12.r
B4499546 lc59c10.y
T49863 ya87f09.r1
T59199 yb50f09.r1
AA746615 nx18d11.s
BF687930 602067010
W10194 ma43h02.r1
BG982434 PMI-CN009
F28020 HSPDL6391 H
AI707460 as90e08.x
BM768418 K-EST0051
BM768693 K-EST0051
AI024403 ov52d04.x
AA407038 EST02022
AA523332 n163c06.s
AI761042 w110g04.x
AA410036 EST01912
AA574047 nk11c03.s
AA644882 v83f12.r
BE620722 601483563
BE833193 QV3-OT006
AI718994 as50h05.x
AA500732 vf96d09.r
AW262820 xq95f02.x
AI153298 uc48b02.r
AI708841 as27f01.x
c 299 18 13.6 376 10 AW884466
300 18 13.6 376 14 D54297
301 18 13.6 376 14 F30615
302 18 13.6 377 9 AL022660
303 18 13.6 377 14 H96057
c 304 18 13.6 378 9 AI084693
305 18 13.6 378 9 AA265400
306 18 13.6 378 9 AW884159
c 307 18 13.6 378 14 BQ549781
308 18 13.6 378 14 AA138321
309 18 13.6 379 12 BF581235
310 18 13.6 379 13 BI522159
c 311 18 13.6 380 9 AA737823
c 312 18 13.6 380 9 AI707617
313 18 13.6 380 14 BM785046
314 18 13.6 381 10 AV663919
315 18 13.6 381 10 AW545053
316 18 13.6 381 10 AW884474
317 18 13.6 381 14 D55058
c 318 18 13.6 382 9 AI000237
c 319 18 13.6 383 9 AI719778
320 18 13.6 383 12 BF808578
c 321 18 13.6 384 9 AA181280
322 18 13.6 384 9 AA559927
323 18 13.6 384 10 AW619405
324 18 13.6 384 12 BE738436
325 18 13.6 385 14 BQ319759
326 18 13.6 386 9 AI206275
327 18 13.6 386 9 AA518100
328 18 13.6 386 14 BM850736
329 18 13.6 386 14 D54613
330 18 13.6 387 9 AA871154
331 18 13.6 387 9 AI911108
332 18 13.6 387 12 BF615106
333 18 13.6 387 13 BM661927
334 18 13.6 387 14 BM855608
335 18 13.6 387 14 BQ784063
336 18 13.6 388 9 AA592421
337 18 13.6 388 14 T63133
338 18 13.6 389 9 AA973422
339 18 13.6 389 9 AI473863
340 18 13.6 389 9 AI830918
341 18 13.6 389 9 AA529989
342 18 13.6 389 10 AW020063
343 18 13.6 390 9 AA849000
344 18 13.6 390 9 AA614380
c 345 18 13.6 390 13 BE18226
346 18 13.6 391 9 AA723043
347 18 13.6 391 9 AI814223
348 18 13.6 391 10 AW163413
349 18 13.6 391 13 BG964620
350 18 13.6 391 14 R31298
351 18 13.6 392 9 AA637703
352 18 13.6 392 9 AI203929
353 18 13.6 392 9 AA083624
354 18 13.6 392 9 AA266123
355 18 13.6 392 14 F24712
356 18 13.6 392 14 D54589
357 18 13.6 393 9 AA915912
358 18 13.6 393 9 AI761022
359 18 13.6 393 9 AA160388
360 18 13.6 393 10 AV697652
361 18 13.6 393 12 BF447868
362 18 13.6 394 9 AA727844
363 18 13.6 394 9 AA871058
364 18 13.6 394 9 AI582139
c 365 18 13.6 394 9 AI707776
366 18 13.6 394 9 AA590481
367 18 13.6 394 10 AW619016
368 18 13.6 395 9 AA649497
369 18 13.6 395 9 AI832679
370 18 13.6 395 10 BE138303
371 18 13.6 395 14 D54570
AW884466 QV3-OT006
D54297 HSPDL35B09B
F30615 HUMD211778 H
AL022660 s7827b40
H96057 yw61906.r1
AI084693 oz79b02.x
AA265400 mz48c05.r
AW884159 QV3-OT006
BQ549781 lk91h08.x
AA138321 mq96c09.r
BF581235 602100449
BI522159 603081565
AA737823 nx14e12.s
AI707617 as31e10.x
BM785046 K-EST0063
AV663919 AV663919
AW545053 C0190B11-
AW884474 QV3-OT006
D55058 HUM163F01B
AI000237 ct04c05.s
AI719778 as41q12.x
BF808578 CM1-CI009
AA181280 zp68d10.s
AA559927 nl60b08.s
AW619405 7456 MARC
BE738436 601572460
BQ319759 PM4-CT081
AI206275 qr22h11.x
AA518100 vl23b05.r
BM850736 K-EST0131
D54613 HUM146A07B
AA871154 vq31q08.r
AI911108 qe18b05.x
BF615106 d881d06.x
BM661927 UI-E-CK1-
BM855608 K-EST0138
BQ784063 SNEST4a83
AA592421 v035d07.r
T63133 yb99a01.s1
AA973422 0044g02.s
AI473863 tal1h03.x
AI830918 wj61h04.x
AA529989 vj13h08.r
AW020063 df03f07.y
AA849000 EST191762
AA614380 nd46c04.s
BE18226 603042056
AA723043 z983a06.s
AI814223 wk54h06.x
AW163413 au83e01.y
BG964620 602831809
R31298 yb76a02.r1
AA637703 vr21b10.r
AI203929 qd79g02.x
AA083624 zn39a02.r
AA266123 mz49a05.r
D54589 HUM145B02B
F24712 HSPDL1231 H
AA915912 oh86d07.s
AI761022 wl10c02.x
AA160388 zo64a07.r
AV697652 AV697652
BF447868 7q96e04.x
AA727844 vp33e12.r
AA871058 vq29g09.r
AI582139 as03c07.x
AI707776 as35d01.x
AA590481 vm24d06.r
AW619016 353 MARC
AA649497 ns44b09.s
AI832679 at71b07.x
BE138303 ug51a05.y
D54570 HUM144D01B

c 372 18 13.6 396 9 A1718322 A1718322 as53b08.x 445 18 13.6 410 9 A816138 A816138 VP44C06.I
c 373 18 13.6 396 9 A4A83076 A4A83076 ne71e11.s 446 18 13.6 410 9 A120749 A120749 mp73c09.I
c 374 18 13.6 396 10 AV762048 AV762048 AV762048 447 18 13.6 410 9 A149000 A149000 to1g11.x
c 375 18 13.6 396 13 B1032602 B1032602 PM3-NN117 448 18 13.6 410 9 A1719722 A1719722 as92b02.x
c 376 18 13.6 396 14 BM991466 BM991466 UI-H-DHO- 449 18 13.6 410 9 A403689 A403689 m255g05.I
c 377 18 13.6 397 9 A157341 A157341 zo42h06.s 450 18 13.6 410 9 A4590198 A4590198 v156a09.I
c 378 18 13.6 398 14 BM687860 BM687860 UI-E-CKI- 451 18 13.6 410 9 A600887 A600887 np56d10.s
c 379 18 13.6 397 9 A1087406 A1087406 o27f11.x 452 18 13.6 410 9 A615404 A615404 v069f02.I
c 380 18 13.6 398 9 A4221676 A4221676 mv28b04.r 453 18 13.6 410 12 BE469210 BE469210 hc79c11.x
c 381 18 13.6 398 9 AA470085 AA470085 xt98g07.s 454 18 13.6 410 12 BE854595 BE854595 ux31f10.Y
c 382 18 13.6 398 10 AV746756 AV746756 AV746756 455 18 13.6 410 13 BM505632 BM505632 lh14h07.x
c 383 18 13.6 398 10 AW619960 AW619960 8046 MARC 456 18 13.6 410 14 BQ129534 BQ129534 i130h11.x
c 384 18 13.6 398 13 B1496662 B1496662 df127b03. 457 18 13.6 411 9 A673311 A673311 v069g12.I
c 385 18 13.6 398 14 BM988447 BM988447 UI-H-DHO- 458 18 13.6 411 9 A870153 A870153 vq12e02.I
c 386 18 13.6 399 14 F24867 F24867 HSPD11526 H 459 18 13.6 411 9 A180646 A180646 at05f11.x
c 387 18 13.6 400 9 AA9793401 AA9793401 vn52e09.r 460 18 13.6 411 9 A208486 A208486 mv85c05.I
c 388 18 13.6 400 9 AA946832 AA946832 qd43e03.s 461 18 13.6 411 9 A527683 A527683 nh86d05.s
c 389 18 13.6 400 9 A1001091 A1001091 e94b03.s 462 18 13.6 411 9 A619410 A619410 v087h12.I
c 390 18 13.6 400 13 B1518034 B1518034 6030411715 463 18 13.6 411 10 AW189058 AW189058 xk99c08.x
c 391 18 13.6 400 14 W05321 W05321 za84e05.r1 464 18 13.6 411 14 BM855519 BM855519 K-EST0138
c 392 18 13.6 401 9 AA864455 AA864455 OH01907.s 465 18 13.6 412 9 AA060514 AA060514 mj68b09.I
c 393 18 13.6 401 9 A1429236 A1429236 ml58d08.x 466 18 13.6 412 9 A871729 A871729 vq40c12.I
c 394 18 13.6 401 9 A1670720 A1670720 wc28d04.x 467 18 13.6 412 9 A1355270 A1355270 qt79e09.x
c 395 18 13.6 401 9 A1872025 A1872025 vm54e11.x 468 18 13.6 412 9 A1924655 A1924655 wa57h09.x
c 396 18 13.6 401 9 AA597004 AA597004 v017h05.I 469 18 13.6 412 10 AW619959 AW619959 8045 MARC
c 397 18 13.6 401 10 AV739443 AV739443 AV739443 470 18 13.6 412 10 AW951791 AW951791 BST363861
c 398 18 13.6 401 10 AW305108 AW305108 xv99f06.x 471 18 13.6 412 12 BG579192 BG579192 df24f01.Y
c 399 18 13.6 401 12 BE833078 BE833078 QV3-OT006 472 18 13.6 413 9 AA734426 AA734426 v190e07.I
c 400 18 13.6 401 14 BM749826 BM749826 K-EST0025 473 18 13.6 413 9 AA929515 AA929515 v140c07.I
c 401 18 13.6 402 9 AA674528 AA674528 vt38f03.I 474 18 13.6 413 9 A1084619 A1084619 o278b10.x
c 402 18 13.6 402 9 AA793944 AA793944 vt38f03.I 475 18 13.6 413 9 A590717 A590717 v059f12.I
c 403 18 13.6 402 9 AA855175 AA855175 aj02c02.s 476 18 13.6 413 10 AV663360 AV663360 8045 MARC
c 404 18 13.6 402 9 A1707686 A1707686 as32a08.x 477 18 13.6 413 12 BG036404 BG036404 602326530
c 405 18 13.6 402 9 A0042227 A0042227 A0042227 478 18 13.6 414 9 AA582794 AA582794 nn35b08.s
c 406 18 13.6 402 9 AA411834 AA411834 zt67f01.s 479 18 13.6 414 13 BI084746 BI084746 602869757
c 407 18 13.6 402 14 BM991448 BM991448 UI-H-DHO- 480 18 13.6 415 9 AA792197 AA792197 v045c06.I
c 408 18 13.6 403 9 AA774770 AA774770 e878f06.s 481 18 13.6 415 9 AA824266 AA824266 aj29e08.s
c 409 18 13.6 403 9 A080144 A080144 mj99f01.I 482 18 13.6 415 9 A870311 A870311 vq14d05.I
c 410 18 13.6 403 9 A1708863 A1708863 as27h03.x 483 18 13.6 415 9 A147960 A147960 zo48d12.s
c 411 18 13.6 403 10 AV760430 AV760430 AV760430 484 18 13.6 415 14 BM850366 BM850366 K-EST0130
c 412 18 13.6 403 10 AW474911 AW474911 xy21e07.x 485 18 13.6 415 14 BM855515 BM855515 K-EST0138
c 413 18 13.6 403 14 BM841476 BM841476 K-EST0118 486 18 13.6 416 9 AA937247 AA937247 oJ02f01.s
c 414 18 13.6 404 9 AA636867 AA636867 vt20g07.I 487 18 13.6 416 9 AA308305 AA308305 EST179135
c 415 18 13.6 404 9 A1718269 A1718269 as51d08.x 488 18 13.6 416 10 BE217903 BE217903 hv31d08.x
c 416 18 13.6 404 9 A1749969 A1749969 ac34e02.x 489 18 13.6 416 10 BE217903 BE217903 hv31d08.x
c 417 18 13.6 404 9 A1832625 A1832625 ac70h07.x 490 18 13.6 416 14 BM856279 BM856279 K-EST0139
c 418 18 13.6 404 9 A1985951 A1985951 wt79b02.x 491 18 13.6 416 14 W91744 W91744 MTA.G01.080
c 419 18 13.6 404 12 BG688032 BG688032 335437 BA 492 18 13.6 417 9 AA755328 AA755328 vq62e09.I
c 420 18 13.6 405 9 AA308256 AA308256 EST179106 493 18 13.6 417 9 A184617 A184617 qf46c03.x
c 421 18 13.6 405 9 AA417155 AA417155 zu04e11.I 494 18 13.6 417 9 A1879910 A1879910 ar99d03.x
c 422 18 13.6 405 9 AA590590 AA590590 vm59f02.I 495 18 13.6 417 9 A1884379 A1884379 vm31f04.x
c 423 18 13.6 405 10 AW189208 AW189208 x102g08.x 496 18 13.6 417 9 AA63174 AA63174 vq85h06.I
c 424 18 13.6 406 9 AA871000 AA871000 vq29a05.I 497 18 13.6 417 9 AA623014 AA623014 np79h12.s
c 425 18 13.6 406 9 AA245555 AA245555 mv50a05.I 498 18 13.6 417 10 AV761674 AV761674 AV761674
c 426 18 13.6 406 12 BG100882 BG100882 uy16e08.Y 499 18 13.6 417 12 BF435945 BF435945 nab75d05.
c 427 18 13.6 407 9 AA823363 AA823363 vp36b10.I 500 18 13.6 417 12 BE906510 BE906510 601498575
c 428 18 13.6 407 9 AA583191 AA583191 nm45b04.s 501 18 13.6 417 12 BF034316 BF034316 601455334
c 429 18 13.6 407 10 AW542557 AW542557 C0159A06- 502 18 13.6 417 13 BI276078 BI276078 UI-R-CW0-
c 430 18 13.6 408 9 A1718723 A1718723 as95e10.x 503 18 13.6 418 9 AA057636 AA057636 z194a08.s
c 431 18 13.6 408 9 AA615456 AA615456 v078e03.I 504 18 13.6 418 9 A846825 A846825 aj99h02.s
c 432 18 13.6 408 10 AW770926 AW770926 h193d10.x 505 18 13.6 418 9 A133875 A133875 qp99d04.x
c 433 18 13.6 408 14 W54335 W54335 md05c05.r1 506 18 13.6 418 9 A1735492 A1735492 at12c08.x
c 434 18 13.6 409 9 AA637951 AA637951 vt22c10.I 507 18 13.6 418 9 A1749009 A1749009 at28h06.x
c 435 18 13.6 409 9 AA665053 AA665053 nu73g12.s 508 18 13.6 418 9 AA402336 AA402336 zt67e12.I
c 436 18 13.6 409 9 A755337 A755337 vq62f09.I 509 18 13.6 419 10 BE629242 BE629242 nu35b06.Y
c 437 18 13.6 409 9 A1333615 A1333615 qp91b04.x 510 18 13.6 419 9 AA758513 AA758513 ah66c02.s
c 438 18 13.6 409 10 AV761739 AV761739 AV761739 511 18 13.6 419 9 AA421628 AA421628 zt98g07.I
c 439 18 13.6 409 10 AW619537 AW619537 7602 MARC 512 18 13.6 419 9 AA506282 AA506282 nh10f05.s
c 440 18 13.6 409 13 BG982683 BG982683 IL5-CN006 513 18 13.6 419 9 AA574728 AA574728 nm31e07.I
c 441 18 13.6 409 14 BM704129 BM704129 UI-E-CKI- 514 18 13.6 419 12 BF036306 BF036306 601459489
c 442 18 13.6 410 9 AA691901 AA691901 vt14g02.I 515 18 13.6 420 9 AA630236 AA630236 ab96f06.s
c 443 18 13.6 410 9 AA770789 AA770789 vt12a11.I 516 18 13.6 420 10 AW619764 AW619764 7831 MARC
c 444 18 13.6 420 10 AW619764 AW619764 7831 MARC 517 18 13.6 420 10 AW619764 AW619764 7831 MARC

518	18	13.6	420	10	AW619801	591	18	13.6	433	13	BM508999
c 519	18	13.6	420	10	AW998270	592	18	13.6	434	9	AA822642
520	18	13.6	420	12	BF426632	593	18	13.6	434	9	AA072373
c 521	18	13.6	421	9	AI142610	594	18	13.6	434	9	AA146886
522	18	13.6	421	9	AA097362	595	18	13.6	434	10	AA464783
c 523	18	13.6	421	10	AV662697	596	18	13.6	434	10	AW990134
524	18	13.6	421	10	AW784283	597	18	13.6	434	12	BE909902
525	18	13.6	421	14	T28934	598	18	13.6	434	14	H71272
c 526	18	13.6	422	9	AA546118	599	18	13.6	434	14	R10426
527	18	13.6	422	14	BM988129	600	18	13.6	434	14	W02899
c 528	18	13.6	423	9	AI190309	601	18	13.6	434	14	W07298
529	18	13.6	423	9	AA445072	602	18	13.6	435	9	AA733430
c 530	18	13.6	423	10	AV592009	603	18	13.6	435	9	AA608354
531	18	13.6	423	12	BF525063	c 604	18	13.6	435	10	AV591223
c 532	18	13.6	424	9	AI718016	605	18	13.6	435	14	BM692911
533	18	13.6	424	13	BM429376	606	18	13.6	435	14	W09121
c 534	18	13.6	424	14	BM987456	607	18	13.6	436	9	AA668938
535	18	13.6	425	9	AA070874	c 608	18	13.6	436	9	AI285228
c 536	18	13.6	425	9	AI085716	609	18	13.6	436	9	AI333853
537	18	13.6	425	9	AI085902	c 610	18	13.6	436	9	AI335740
c 538	18	13.6	425	9	AI192730	611	18	13.6	436	9	AA208203
539	18	13.6	425	9	AA081417	612	18	13.6	436	10	AV760449
c 540	18	13.6	425	10	AV711964	613	18	13.6	436	10	AW619007
541	18	13.6	425	10	AW619595	614	18	13.6	436	10	BE226076
c 542	18	13.6	426	9	AI282536	615	18	13.6	436	12	BF724548
543	18	13.6	426	14	BM853886	616	18	13.6	436	12	BE855183
c 544	18	13.6	427	9	AA961747	617	18	13.6	436	13	BI186763
545	18	13.6	427	9	AA098544	618	18	13.6	436	14	BM735404
c 546	18	13.6	427	12	BE751572	619	18	13.6	436	14	BM796835
547	18	13.6	427	13	BI848878	c 620	18	13.6	436	14	N70669
c 548	18	13.6	427	13	BM125857	621	18	13.6	436	14	W12431
549	18	13.6	427	14	BM768493	c 622	18	13.6	437	9	AI184003
c 550	18	13.6	428	9	AA314732	623	18	13.6	437	9	AA305206
551	18	13.6	428	9	AA314732	624	18	13.6	437	9	AA624275
c 552	18	13.6	428	9	AA483378	625	18	13.6	437	10	AW238493
553	18	13.6	428	10	AA463456	626	18	13.6	437	10	AW465237
c 554	18	13.6	428	10	AA465926	627	18	13.6	437	10	AW619688
555	18	13.6	428	12	BF452590	628	18	13.6	437	10	AW619791
c 556	18	13.6	428	12	BE867472	629	18	13.6	437	14	BQ051557
557	18	13.6	428	12	BE907243	630	18	13.6	438	9	AA714158
c 558	18	13.6	428	13	BM055733	631	18	13.6	438	9	AI952657
559	18	13.6	429	9	AA261583	c 632	18	13.6	438	12	BE710677
c 560	18	13.6	429	10	AW239999	633	18	13.6	438	13	BM669065
561	18	13.6	429	12	BF715155	634	18	13.6	438	14	W36426
c 562	18	13.6	429	12	BF467061	635	18	13.6	439	9	AA048607
563	18	13.6	429	12	BE710616	636	18	13.6	439	9	AA069942
c 564	18	13.6	430	10	AW912908	637	18	13.6	439	9	AA070781
565	18	13.6	430	14	W07728	c 638	18	13.6	439	9	AI150260
c 566	18	13.6	431	9	AI346709	639	18	13.6	439	9	AA146782
567	18	13.6	431	10	AW619277	640	18	13.6	439	9	AA610379
c 568	18	13.6	431	12	BG253202	641	18	13.6	439	9	AA631486
569	18	13.6	431	12	BE889330	642	18	13.6	439	12	BG422271
c 570	18	13.6	431	14	N70636	643	18	13.6	439	12	BF043545
571	18	13.6	431	14	R28160	644	18	13.6	439	14	BM705341
c 572	18	13.6	432	9	AA059707	645	18	13.6	439	17	AQ337608
573	18	13.6	432	9	AA633937	c 646	18	13.6	440	9	AI120438
c 574	18	13.6	432	9	AA082482	647	18	13.6	440	9	AI178899
575	18	13.6	432	9	AI720365	648	18	13.6	440	9	AA518877
c 576	18	13.6	432	9	AA188050	649	18	13.6	440	12	BE909639
577	18	13.6	432	9	AA316182	650	18	13.6	440	12	BE909639
c 578	18	13.6	432	9	AA529971	651	18	13.6	440	14	BI304719
579	18	13.6	432	10	AV618094	652	18	13.6	441	9	BQ637482
c 580	18	13.6	432	10	AV618095	653	18	13.6	441	9	AA050290
581	18	13.6	432	10	AV664771	c 654	18	13.6	441	9	AI051791
c 582	18	13.6	432	10	AV712821	655	18	13.6	441	9	AA106001
583	18	13.6	432	10	BE254902	656	18	13.6	441	9	AA147288
c 584	18	13.6	432	14	BM738252	657	18	13.6	441	9	AA314292
585	18	13.6	432	14	BM738720	658	18	13.6	441	9	AA616372
c 586	18	13.6	433	9	AA730441	659	18	13.6	441	10	BE484103
587	18	13.6	433	9	AI740739	660	18	13.6	441	12	BE875920
c 588	18	13.6	433	9	AA484097	661	18	13.6	441	14	BQ030697
589	18	13.6	433	12	BG223174	c 662	18	13.6	441	14	BQ044096
c 590	18	13.6	433	13	BI304486	663	18	13.6	441	14	W52576

664 18 13.6 442 9 AA929872 v244c07.r 737 18 13.6 450 9 AA148023
665 18 13.6 442 9 A1833129 at75804.x 738 18 13.6 450 9 AA186647
666 18 13.6 442 9 AL022773 87c19p49 739 18 13.6 450 9 AA557805
667 18 13.6 442 9 AA155816 2047d02.r 740 18 13.6 450 10 BE613621
668 18 13.6 442 9 AA230554 mv73e12.r 741 18 13.6 450 14 BM888223
669 18 13.6 442 9 AA420470 nc60f03.r 742 18 13.6 450 14 BQ717613
670 18 13.6 442 9 AA590434 vm20b07.r 743 18 13.6 451 9 AA574701
671 18 13.6 442 9 AA617744 np51c07.s 744 18 13.6 451 12 BG253383
672 18 13.6 442 9 AA623351 vp25b12.r 745 18 13.6 451 13 BM089275
673 18 13.6 442 10 AV615168 AV615168 746 18 13.6 451 14 W04856
674 18 13.6 442 10 AW986772 uf87e06.y 747 18 13.6 452 9 AA793069
675 18 13.6 442 12 BG249640 602319733 748 18 13.6 452 9 A1679886
676 18 13.6 442 13 BM105914 509497 MA 749 18 13.6 452 9 AA480527
677 18 13.6 443 9 AA691560 vs16d07.r 750 18 13.6 452 10 AW989988
678 18 13.6 443 9 A1015078 ov51e04.s 751 18 13.6 452 12 BF614970
679 18 13.6 443 10 BE626900 uu04f08.y 752 18 13.6 452 13 BM431924
680 18 13.6 443 12 BF046634 BP250020A 753 18 13.6 452 14 BQ044342
681 18 13.6 443 14 BQ129779 1130g11.y 754 18 13.6 452 14 C88595
682 18 13.6 443 14 BQ234159 hd41h12.g 755 18 13.6 453 13 BM117611
683 18 13.6 443 14 BQ925777 AGENCOURT 756 18 13.6 453 13 BM431970
684 18 13.6 444 9 AA472840 v458f09.r 757 18 13.6 453 14 BM970898
685 18 13.6 444 9 AA557716 n147b06.s 758 18 13.6 453 14 BQ220112
686 18 13.6 444 9 AA597166 vc35h05.r 759 18 13.6 454 9 AA675430
687 18 13.6 444 10 AV664309 AV664309 760 18 13.6 454 9 A1103001
688 18 13.6 444 10 AW645087 BP230017B 761 18 13.6 454 9 AAG24249
689 18 13.6 444 12 BG671616 DRNBVE09 762 18 13.6 454 12 BQ687948
690 18 13.6 444 14 BM969729 UI-CF-DU1 763 18 13.6 454 14 BQ050657
691 18 13.6 444 14 BQ221715 AGENCOURT 764 18 13.6 455 9 AA869073
692 18 13.6 445 14 BQ563782 g106d04.y 765 18 13.6 455 9 A1049606
693 18 13.6 445 9 AA104455 mc45c09.r 766 18 13.6 455 9 A1254343
694 18 13.6 445 9 A1707865 as26a08.x 767 18 13.6 455 9 A129780
695 18 13.6 445 9 AA207529 mv89c06.r 768 18 13.6 455 9 AAG24360
696 18 13.6 445 9 AA208172 mv82d11.r 769 18 13.6 455 10 AV762708
697 18 13.6 445 10 AW669799 113369 MA 770 18 13.6 455 10 AW082906
698 18 13.6 445 14 BM841330 K-EST0118 771 18 13.6 455 10 AW619290
699 18 13.6 445 14 N99549 za40b01.r1 772 18 13.6 455 10 AW917430
700 18 13.6 446 9 AA104283 mo45e09.r 773 18 13.6 455 10 BE113191
701 18 13.6 446 9 A1719522 as44d03.x 774 18 13.6 455 12 BE908679
702 18 13.6 446 9 AA245919 mx06g06.r 775 18 13.6 455 12 BE967731
703 18 13.6 446 9 AA240699 nc62d07.r 776 18 13.6 455 14 BM763917
704 18 13.6 446 10 BE273236 601142546 777 18 13.6 455 14 BQ600351
705 18 13.6 446 12 BGL68277 602340814 778 18 13.6 456 9 AA659067
706 18 13.6 446 13 B1327377 AR077E02T 779 18 13.6 456 9 AA437971
707 18 13.6 446 14 BM760559 K-EST0041 780 18 13.6 456 10 AW662449
708 18 13.6 446 14 BQ028237 UI-H-COO- 781 18 13.6 456 10 AW619874
709 18 13.6 446 14 R11502 yf28f11.s1 782 18 13.6 456 12 BF339067
710 18 13.6 446 14 T59161 yb50f09.s1 783 18 13.6 456 13 BM542700
711 18 13.6 446 14 W37054 mb72h10.r1 784 18 13.6 456 14 BM766933
712 18 13.6 446 14 W44429 zc28g02.s1 785 18 13.6 456 14 BQ085873
713 18 13.6 447 9 A1200768 qf67b10.x 786 18 13.6 456 14 W29255
714 18 13.6 447 9 A1749605 at25b06.x 787 18 13.6 456 14 W29917
715 18 13.6 447 9 AAG22812 np77c11.s 788 18 13.6 457 9 AA665306
716 18 13.6 447 14 BM738093 K-EST0002 789 18 13.6 457 9 AA538207
717 18 13.6 447 14 BM888592 FMW271 Hu 790 18 13.6 457 10 AV663920
718 18 13.6 448 9 AA844493 aj21b03.s 791 18 13.6 457 12 BGL50602
719 18 13.6 448 9 A1313428 qp80d09.x 792 18 13.6 457 12 BG778264
720 18 13.6 448 9 AA417156 zu04e12.r 793 18 13.6 457 13 B1521153
721 18 13.6 448 10 AW161076 au78f08.y 794 18 13.6 457 13 B1838135
722 18 13.6 448 13 BM424102 AGENCOURT 795 18 13.6 457 14 BM818373
723 18 13.6 448 14 W61175 zd31b01.r1 796 18 13.6 457 14 BQ234689
724 18 13.6 449 9 A1327405 mq37c10.x 797 18 13.6 458 9 AA822935
725 18 13.6 449 12 BF034546 601455457 798 18 13.6 458 9 A1249439
726 18 13.6 449 13 B1848716 471229 MA 799 18 13.6 458 9 AAG60682
727 18 13.6 449 13 B1848728 471246 MA 800 18 13.6 458 12 BG506360
728 18 13.6 449 13 BM015450 603641643 801 18 13.6 459 9 A1180916
729 18 13.6 449 13 BM088974 502602 MA 802 18 13.6 459 9 A1679377
730 18 13.6 449 14 BQ674882 AGENCOURT 803 18 13.6 459 9 AA409574
731 18 13.6 449 14 W08781 mb51e12.r1 804 18 13.6 459 12 BF526927
732 18 13.6 449 14 W09113 ma70c04.r1 805 18 13.6 459 12 BF276605
733 18 13.6 450 9 AA869790 vq15h03.r 806 18 13.6 459 12 BG059103
734 18 13.6 450 9 AA881962 vx30h08.r 807 18 13.6 459 12 BG113182
735 18 13.6 450 9 A1000052 os50a03.s 808 18 13.6 459 12 BE904152
736 18 13.6 450 9 A1081185 or32d11.x 809 18 13.6 459 12 BE904152

AA148023 z048d12.r
AA186647 zp04g03.r
AA557805 nl62g09.s
BE613621 601504410
BM888223 TMM137 Hu
BQ717613 AGENCOURT
AA574701 vm31d09.r
BG253383 602362987
BG961429 MR4-CT053
BM089275 502986 MA
W04856 za75a12.r1
AA793069 vp21c12.r
A1679886 tu67a08.x
AA480527 ne63b09.s
AA989988 uf31g08.y
BF614970 dg80c03.x
BM431924 1JEJ14B5
BQ044342 UI-M-EF0-
C88595 C88595 Mous
BM117611 L0853H05-
BM431970 1JEJ14G2
BM970898 UI-CF-EC1
BQ220112 AGENCOURT
AA675430 v101c07.r
A1103001 EST212290
AAG24249 vm98e10.r
BQ687948 335289 BA
BQ050657 AGENCOURT
AA869073 vq30d08.r
A1049606 an38b03.x
A1254343 qn5b10.x
AA129780 zu51a04.r
AAG24360 vm20e02.r
AV762708 AV762708
AW082906 xc04c01.x
AW619290 652 MARC
AW917430 EST348734
BE113191 UI-R-BJ1-
BE908679 601499520
BE967731 601648704
BM763917 K-EST0045
BQ600351 MI-P-E7-a
BQ659067 nu81d01.s
AA437971 v332h08.r
AA662449 AV662449
AW619874 7954 MARC
BF339067 602034860
BM542700 AGENCOURT
BM766933 K-EST0048
BQ085873 i111b07.y
W29255 mc18g02.r1
W29917 mc08c04.r1
AA665306 nt88g03.s
AA538207 v156h08.r
AA663920 AV663920
BGL50602 na132c09.
BG778264 602666852
B1521153 603081565
B1838135 603083637
BM818373 K-EST0085
BQ234689 hd48g02.g
AA822935 vp30f10.r
A1249439 qx04e11.x
AA606682 vm93g02.r
BG506360 601860425
A1180916 ub77c03.r
A1679377 tu63a08.x
AA409574 EST01315
BF526927 602070201
BF276605 by09b10.y
BG059103 nah45b12.
BG113182 602283853
BE904152 601494365

C 810	18	13.6	459	13	B1281171	UI-R-CT0s	B1281171	UI-R-CT0s	C 883	18	13.6	469	14	BM994578	UI-H-DH0-
C 811	18	13.6	459	14	B0668952	AGENCY	B0668952	AGENCY	884	18	13.6	469	14	BQ050650	AGENCY
C 812	18	13.6	459	14	BQ669786	AGENCY	BQ669786	AGENCY	885	18	13.6	470	14	BM840001	K-EST0111
C 813	18	13.6	460	9	A1057027	WY75D02.x	A1057027	WY75D02.x	886	18	13.6	470	14	BM842129	K-EST0119
C 814	18	13.6	460	10	AW044191	WY71C02.x	AW044191	WY71C02.x	887	18	13.6	470	14	BQ222029	AGENCY
C 815	18	13.6	460	12	BG807231	2052-81 M	BG807231	2052-81 M	C 888	18	13.6	471	9	AA420669	nc52d06.s
C 816	18	13.6	460	12	BF021823	uy58e10.y	BF021823	uy58e10.y	C 889	18	13.6	471	9	AA492043	ng14c06.s
C 817	18	13.6	460	14	BM806254	AGENCY	BM806254	AGENCY	C 890	18	13.6	471	9	AA493908	nh07d02.s
C 818	18	13.6	460	14	BQ230686	AGENCY	BQ230686	AGENCY	891	18	13.6	471	10	AV591224	AV591224
C 819	18	13.6	461	9	A1024383	ov52a10.x	A1024383	ov52a10.x	C 892	18	13.6	471	12	BG376039	UI-R-CS0-
C 820	18	13.6	461	9	AA510551	vh59g02.r	AA510551	vh59g02.r	893	18	13.6	471	13	B1187094	AR048A05L
C 821	18	13.6	461	10	AW913388	uf52d09.y	AW913388	uf52d09.y	894	18	13.6	472	10	BEG26403	uu05408.y
C 822	18	13.6	461	12	BF012437	ux56c09.y	BF012437	ux56c09.y	895	18	13.6	472	14	BM840163	K-EST0117
C 823	18	13.6	461	14	BM739041	K-EST0008	BM739041	K-EST0008	896	18	13.6	472	14	BM840163	AGENCY
C 824	18	13.6	461	14	BQ937688	AGENCY	BQ937688	AGENCY	C 897	18	13.6	472	14	BQ019521	UI-H-ED0-
C 825	18	13.6	461	14	W44487	zg28g02.r1	W44487	zg28g02.r1	C 898	18	13.6	473	13	BM210025	C0657H03-
C 826	18	13.6	462	9	AA407511	EST00785	AA407511	EST00785	899	18	13.6	473	14	BQ044034	UI-M-EF0-
C 827	18	13.6	462	10	AW913388	uf52d09.y	AW913388	uf52d09.y	900	18	13.6	473	14	BQ050822	AGENCY
C 828	18	13.6	462	12	BF972425	602240588	BF972425	602240588	901	18	13.6	473	14	BQ570076	g1144a05
C 829	18	13.6	462	13	BM471265	AGENCY	BM471265	AGENCY	902	18	13.6	474	11	AK012361	Mus muscu
C 830	18	13.6	463	9	AA684524	vr02d03.r	AA684524	vr02d03.r	903	18	13.6	474	13	BM460102	AGENCY
C 831	18	13.6	463	9	AA985724	ue13e02.y	AA985724	ue13e02.y	904	18	13.6	474	14	BQ215042	AGENCY
C 832	18	13.6	463	10	AW664772	AV664772	AW664772	AV664772	905	18	13.6	475	9	A1035790	ue17c09.y
C 833	18	13.6	463	10	AW544206	C0179A04-	AW544206	C0179A04-	906	18	13.6	475	9	A1157351	ui45f09.y
C 834	18	13.6	463	10	AW546474	L0008G02-	AW546474	L0008G02-	C 907	18	13.6	475	9	AA525997	ni58b07.s
C 835	18	13.6	463	10	BE117292	UI-R-BS1-	BE117292	UI-R-BS1-	908	18	13.6	475	11	AK005512	Mus muscu
C 836	18	13.6	463	10	BE680811	df85e07.y	BE680811	df85e07.y	C 909	18	13.6	475	13	BI285116	UI-R-DB0-
C 837	18	13.6	463	12	BF033570	G01453537	BF033570	G01453537	910	18	13.6	475	13	BI17803	603042056
C 838	18	13.6	463	13	B1278755	UI-R-C20-	B1278755	UI-R-C20-	911	18	13.6	475	14	W12205	ma65c12.r1
C 839	18	13.6	463	14	BQ672252	AGENCY	BQ672252	AGENCY	C 912	18	13.6	475	17	AQ974928	RPCI-23-3
C 840	18	13.6	464	9	AA207582	mv81a10.r	AA207582	mv81a10.r	913	18	13.6	476	9	AA409035	EST03526
C 841	18	13.6	464	9	AA208305	mv82f08.r	AA208305	mv82f08.r	914	18	13.6	476	11	AK008557	Mus muscu
C 842	18	13.6	464	9	AA579241	nf36c12.s	AA579241	nf36c12.s	915	18	13.6	476	11	AK019212	Mus muscu
C 843	18	13.6	464	12	BG252807	G02365375	BG252807	G02365375	C 916	18	13.6	476	12	BG063081	H3001H10-
C 844	18	13.6	464	14	BM668226	UI-E-CK1-	BM668226	UI-E-CK1-	917	18	13.6	476	14	BM915571	AGENCY
C 845	18	13.6	464	14	BM739931	K-EST0010	BM739931	K-EST0010	918	18	13.6	477	10	BE1138152	u349D05.y
C 846	18	13.6	464	14	BQ231135	AGENCY	BQ231135	AGENCY	919	18	13.6	477	11	AK019232	Mus muscu
C 847	18	13.6	464	14	W16623	zbo5g11.r1	W16623	zbo5g11.r1	920	18	13.6	477	12	BG114175	G02286002
C 848	18	13.6	464	14	W85257	mf49g12.r1	W85257	mf49g12.r1	921	18	13.6	477	12	BG119390	G02349073
C 849	18	13.6	465	9	A1912423	tz05d05.x	A1912423	tz05d05.x	C 922	18	13.6	477	14	BM990256	UI-H-DH0-
C 850	18	13.6	465	9	AU122952	AU122952	AU122952	AU122952	C 923	18	13.6	477	14	BM990256	UI-H-DH0-
C 851	18	13.6	465	9	AA600818	np45p04.s	AA600818	np45p04.s	C 924	18	13.6	477	14	BQ564389	g115g04.y
C 852	18	13.6	465	13	BM423572	AGENCY	BM423572	AGENCY	C 925	18	13.6	477	14	BQ601337	MT-P-H3-a
C 853	18	13.6	465	14	BQ712482	AGENCY	BQ712482	AGENCY	C 926	18	13.6	478	9	AA397446	nc65e06.r
C 854	18	13.6	466	9	AA956716	UI-E-E1-f	AA956716	UI-E-E1-f	927	18	13.6	478	10	AA589574	AV589574
C 855	18	13.6	466	13	BM432631	LJEJ12A8.	BM432631	LJEJ12A8.	928	18	13.6	478	12	BG028891	G02292392
C 856	18	13.6	466	14	BQ222149	AGENCY	BQ222149	AGENCY	C 929	18	13.6	478	13	BM553481	AGENCY
C 857	18	13.6	466	14	BQ569297	g1125h07.	BQ569297	g1125h07.	930	18	13.6	478	14	BM800103	AGENCY
C 858	18	13.6	466	14	BQ675833	AGENCY	BQ675833	AGENCY	931	18	13.6	479	12	BG110782	AGENCY
C 859	18	13.6	467	9	A1019390	ub23c11.r	A1019390	ub23c11.r	932	18	13.6	479	12	BG800713	G02279091
C 860	18	13.6	467	10	AV590225	AV590225	AV590225	AV590225	933	18	13.6	479	12	BG800713	G02279091
C 861	18	13.6	467	13	BI849526	472007 MA	BI849526	472007 MA	934	18	13.6	479	13	BG964407	G02832048
C 862	18	13.6	467	13	BI985272	3133-62 M	BI985272	3133-62 M	C 935	18	13.6	479	14	BQ044293	UI-M-EF0-
C 863	18	13.6	467	14	BM768704	K-EST0051	BM768704	K-EST0051	C 936	18	13.6	479	14	BQ234389	h444g03.g
C 864	18	13.6	467	14	BM916304	AGENCY	BM916304	AGENCY	C 937	18	13.6	479	14	BQ598859	MT-P-E4-a
C 865	18	13.6	467	14	BM988402	UI-H-DH0-	BM988402	UI-H-DH0-	C 938	18	13.6	479	14	C78912	Mous
C 866	18	13.6	467	14	R28159	yh65c09.r1	R28159	yh65c09.r1	939	18	13.6	479	14	W56104	zc58h01.r1
C 867	18	13.6	468	9	A1322354	mc48h10.y	A1322354	mc48h10.y	940	18	13.6	479	14	W56488	zc59c08.r1
C 868	18	13.6	468	14	BM818021	UI-H-DH0-	BM818021	UI-H-DH0-	941	18	13.6	481	9	AU134685	AU134685
C 869	18	13.6	468	14	BM988472	UI-H-DH0-	BM988472	UI-H-DH0-	C 942	18	13.6	481	9	AA537096	v196c09.r
C 870	18	13.6	468	14	BM994288	AGENCY	BM994288	AGENCY	943	18	13.6	481	13	BI285295	UI-R-DB0-
C 871	18	13.6	469	9	AA117808	m911b16	AA117808	m911b16	944	18	13.6	481	14	BQ043863	UI-M-EF0-
C 872	18	13.6	469	9	AU122997	AGENCY	AU122997	AGENCY	945	18	13.6	482	10	AV713635	AV713635
C 873	18	13.6	469	9	AU126196	AU126196	AU126196	AU126196	946	18	13.6	482	11	AK003178	Mus muscu
C 874	18	13.6	469	9	AA533244	rj66a02.s	AA533244	rj66a02.s	947	18	13.6	482	12	BF129865	601817658
C 875	18	13.6	469	10	AV733721	AV733721	AV733721	AV733721	948	18	13.6	482	13	BM453440	AGENCY
C 876	18	13.6	469	10	AM652975	101444 MA	AM652975	101444 MA	949	18	13.6	482	14	BQ220027	AGENCY
C 877	18	13.6	469	13	BI756838	603029024	BI756838	603029024	950	18	13.6	482	14	BQ220027	AGENCY
C 878	18	13.6	469	13	BM114521	L0806A08-	BM114521	L0806A08-	C 951	18	13.6	483	13	BI403636	MT-P-CP1-
C 879	18	13.6	469	14	BM974640	UI-CF-EC1	BM974640	UI-CF-EC1	952	18	13.6	483	13	BM552813	AGENCY
C 880	18	13.6	469	14	BM993626	UI-H-DH0-	BM993626	UI-H-DH0-	C 953	18	13.6	484	12	BF708578	MI-P-AY0-
C 881	18	13.6	469	14	BM994413	UI-H-DH0-	BM994413	UI-H-DH0-	954	18	13.6	484	12	BF427169	df63d01.y
C 882	18	13.6	469	14	BM994413	UI-H-DH0-	BM994413	UI-H-DH0-	C 955	18	13.6	485	13	BI186352	UNL-P-FN-

956 18 13.6 485 14 BM904542
 957 18 13.6 486 12 BG029162
 958 18 13.6 486 14 BM925833
 959 18 13.6 486 14 BQ044617
 960 18 13.6 486 14 BQ276900
 961 18 13.6 487 9 AA924288
 962 18 13.6 487 10 BE135516
 963 18 13.6 487 13 BM552487
 964 18 13.6 487 14 BQ015863
 965 18 13.6 487 14 BQ214887
 966 18 13.6 487 14 BQ688840
 967 18 13.6 488 13 BI911020
 968 18 13.6 488 14 BM920153
 969 18 13.6 488 14 BQ044277
 970 18 13.6 489 13 BI907023
 971 18 13.6 489 14 BQ043931
 972 18 13.6 490 12 BF392651
 973 18 13.6 491 9 AV133286
 974 18 13.6 491 14 BQ044155
 975 18 13.6 492 13 BI561941
 976 18 13.6 493 12 BF652397
 977 18 13.6 493 12 BG715095
 978 18 13.6 493 12 BG777915
 979 18 13.6 493 13 BI602996
 980 18 13.6 493 14 BM905815
 981 18 13.6 494 13 BI494623
 982 18 13.6 494 14 BM951454
 983 18 13.6 495 10 AV758795
 984 18 13.6 495 12 BF977790
 985 18 13.6 495 12 BG705343
 986 18 13.6 495 12 BG715004
 987 18 13.6 495 12 BF185700
 988 18 13.6 495 13 BI597542
 989 18 13.6 495 14 BM990169
 990 18 13.6 495 14 BM997291
 991 18 13.6 495 14 BQ434624
 992 18 13.6 496 9 AI386131
 993 18 13.6 496 12 BF436663
 994 18 13.6 496 12 BF688118
 995 18 13.6 496 13 BI602447
 996 18 13.6 496 13 BI669025
 997 18 13.6 496 14 BM989712
 998 18 13.6 497 10 BF369357
 999 18 13.6 497 12 BF439312
 1000 18 13.6 497 12 BG713987

ALIGNMENTS

RESULT 1
 AW620090 374 bp mRNA linear EST 24-MAR-2000
 LOCUS 852 MARC PBE Sus scrofa cDNA 5', mRNA sequence.
 DEFINITION AW620090
 ACCESSION AW620090.1 GI:7326274
 VERSION EST.
 KEYWORDS pig.
 SOURCE Sus scrofa
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
 1 (bases 1 to 374)
 Smith, T.P.L., Fahrenkrug, S.C., Rohrer, G.A., Simmen, F.A., Rexroad
 , C.E. and Keele, J.W.
 Mapping of expressed sequence tags from a porcine early embryonic
 cDNA library
 Anim. Genet. 32 (2), 66-72 (2001)
 21314990
 Contact: Smith TPL
 USDA, ARS, US Meat Animal Research Center
 PO Box 166, Clay Center, NE 68933-0166, USA
 Tel: 402 762 4366
 Fax: 402 762 4390

BM904542 AGENCOURT
 BG029162 602295257
 BM925833 AGENCOURT
 BQ044617 UI-M-EFO-
 BQ276900 AGENCOURT
 AA924288 UI-R-AI-d
 BE135516 ug52a06.Y
 BM552487 AGENCOURT
 BQ015863 UI-H-DT1-
 BQ214887 AGENCOURT
 BQ688840 AGENCOURT
 BI911020 603068742
 BM920153 AGENCOURT
 BQ044277 UI-M-EFO-
 BI907023 603064979
 BQ043931 UI-M-EFO-
 BF392651 UI-R-CAO-
 AV133286 AV133286
 BQ044155 UI-M-EFO-
 BI561941 603254859
 BF652397 275944 NA
 BG715095 602676033
 BG777915 60265666
 BI602996 603251201
 BM905815 AGENCOURT
 BI494623 dfl111f09.
 BM951454 UI-M-EFO-
 AV758795 AV758795
 BF977790 602148622
 BG705343 602687862
 BG715004 602673209
 BF185700 601817273
 BI597542 603247071
 BM990169 UI-H-DHO-
 BM997291 UI-H-DHO-
 BQ434624 AGENCOURT
 AI386131 mq37c10.Y
 BF436663 7p15h09.x
 BF688118 602067344
 BI602447 603252238
 BI669025 603295120
 BM989712 UI-H-DHO-
 BE369357 601220724
 BF439312 nab62g08.
 BG713987 602674389

Email: smith@email.marc.usda.gov
 Single pass sequencing. Bases called and alt_trimmed with phred
 v0.980904.e. Vector identified by cross_match with the -minscore 18
 and -minmatch 12 options.
 PCR Primers
 FORWARD: GGAACAGCTATGACCAATG
 BACKWARD: GTAAACGACGGCCAGT
 Seq primer: AATTAACCTCACTAAAGG.
 Location/Qualifiers
 1..374
 /organism="Sus scrofa"
 /db_xref="taxon:9823"
 /clone_lib="MARC PBE"
 /tissue_type="Day 12 whole embryos"
 /lab_host="XL0LR"
 /note="Vector: pBLUESCRIPT SK-; Site.1: EcoRI; Site.2:
 XhoI; Library made from pool of embryos in spherical and
 filamentous stages of development (.75% and 92.5%, et al,
 respectively, of each stage) as described in Choi et al,
 Endocrinology 137, 1457-67, 1996."
 BASE COUNT 112 a 86 c 118 g 58 t
 ORIGIN
 Query Match 18.9%; Score 25; DB 10; Length 374;
 Best Local Similarity 100.0%; Pred. No. 0.027; 0; Indels 0; Gaps 0;
 Matches 25; Conservative 0; Mismatches 0;
 QY 33 CAGCTTGGATAAGGCCAAGCTGAAG 57
 Db 98 CAGCTTGGATAAGGCCAAGCTGAAG 122
 RESULT 2
 BG410966 382 bp mRNA linear EST 13-MAR-2001
 LOCUS EML_28_B04_g1_A002 Embryo 1 (EML) Sorghum bicolor cDNA, mRNA
 DEFINITION BG410966
 ACCESSION BG410966.1 GI:13316519
 VERSION EST.
 KEYWORDS sorghum.
 SOURCE Sorghum bicolor
 ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
 clade; Panicoideae; Andropogoneae; Sorghum.
 1 (bases 1 to 382)
 Reid, S.P., Cordonnier-Pratt, M.-M., Gingle, A. and Pratt, L.H.
 An EST database from Sorghum: developing embryos
 Unpublished (2000)
 Contact: Cordonnier-Pratt MM
 Laboratory for Genomics and Bioinformatics
 The University of Georgia, Department of Plant Biology
 Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
 Tel: 706 542 1860
 Fax: 706 583 0210
 Email: mmpratt@uga.edu
 Sequences have been trimmed to exclude PolyA, vector and regions
 below Phred quality 16. The threshold for highest quality sequence
 is 20.
 Seq primer: PolyTMix
 High quality sequence start: 75
 High quality sequence stop: 379
 POLYA=Yes.
 Location/Qualifiers
 1..382
 /organism="Sorghum bicolor"
 /db_xref="taxon:4558"
 /clone_lib="Embryo 1 (EML)"
 /note="Organ: Embryos germinated for 24 hr; Vector:
 pBluescript II from Lambda Zap II; Site.1: XhoI; Site.2:
 EcoRI; The library was made from poly-A RNA in the cloning
 vector lambda Zap II. Clones to be sequenced were
 prepared by mass excision."

BASE COUNT 133 a 77 c 94 g 78 t
ORIGIN

Query Match 15.9%; Score 21; DB 12; Length 382;
Best Local Similarity 100.0%; Pred. No. 3.2;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 GACCTGGAGAAATGCCAGC 36
|||||
Db 227 GACCTGGAGAAATGCCAGC 247

RESULT 3

BMI25588/c

LOCUS

DEFINITION 418 bp mRNA linear EST 12-MAR-2002
musculus cDNA clone IMAGE:5668956 3' similar to SW:TYB0_HUMAN
P13472 THYNOSIN BETA-10. ; mRNA sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 418)

REFERENCE

AUTHORS

Melton, D., Brown, J., Kenty, G., Permutt, A., Lee, C., Kaestner, K.,
Lemishka, I., Scearce, M., Brestelli, J., Gradwohl, G., Clifton, S.,
Hillier, L., Marra, M., Pape, D., Wylie, T., Martin, J., Blistain, A.,
Schmitt, A., Theising, B., Ritter, E., Ronko, I., Bennett, J., Cardenas
, M., Gibbons, M., McCann, R., Cole, R., Tsagareishvili, R., Williams, T.,
Jackson, Y., and Bowers, Y.
Endocrine Pancreas Consortium
Unpublished (2000)

TITLE

JOURNAL

COMMENT

Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
Endocrine Pancreas Consortium
Harvard University, Howard Hughes Medical Institute
Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
MA 02138

Tel: 617-495-1812

Fax: 617-495-8557

Email: dmelton@biohp.harvard.edu

Library was constructed by Dr. Douglas Melton DNA sequencing by:
Washington University Genome Sequencing Center For information on
obtaining a clone please contact: Juliana Brown
(brown@fas.harvard.edu)

MGI:1955282 This sequence now available from the IMAGE consortium,
for clone orders contact: info@image.llnl.gov
High quality sequence stop: 66.

FEATURES

source

1. .418

Location/Qualifiers

/organism="Mus musculus"

/strain="ICR"

/db_xref="taxon:10090"

/clone="IMAGE:5668956"

/clone_lib="Melton Normalized Mixed Mouse Pancreas 1

NI-MMS1"

/sex="Both for embryonic & newborn, male for adult and

adult islet"

/dev_stage="Embryonic day 10.5, E12.5, E16.5, newborn,

adult, mixed"

/lab_host="DH10B"

/note="Vector: pSPORT1; Site.1: Not I; Site.2: Sal I; Five
libraries representing E10.5/12.5 pancreatic bud, E16.5
pancreas, newborn pancreas, adult pancreas, and adult
islets of Langerhans were separately constructed using
Superscript Plasmid Library kit (Life Technologies). cDNA
was made by oligo-dr priming and size-selected by column
fractionation. Libraries were amplified once on solid
support and plasmid DNA from each library was prepared
and mixed in equal amounts. The mixed library DNA was
normalized by method #4 from Bonaldo, Lennon, and Soares
1996 Genome Research 6:791-806; 0.5 microgram

single-stranded mixed library plasmid DNA was mixed with
5 micrograms PCR product representing mixed library
inserts and hybridized to an EcoT of 6. Single-stranded
(unhybridized) plasmids were isolated by hydroxyapatite
chromatography and used to make this library."

BASE COUNT 68 a 98 c 116 g 133 t
ORIGIN

Query Match 15.9%; Score 21; DB 13; Length 418;
Best Local Similarity 100.0%; Pred. No. 3.3;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 37 TTGGATAAGCCCAAGCTGAAG 57
|||||
Db 385 TTGGATAAGCCCAAGCTGAAG 365

RESULT 4

AZ098162

LOCUS

DEFINITION 608 bp DNA linear GSS 09-MAY-2001
RPCI-23-15E12-TV RPCI-23 Mus musculus genomic clone RPCI-23-15E12,
DNA sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Mus musculus
house mouse.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 608)
Zhao, S., Nierman, W., Feldblyum, T., Malek, J., Shatsman, S., Akınret
, B., Levins, M., McGann, S., Tsegaye, G., Geer, K., Krol, M., de Jong, P.
and Fraser, C.M.
Mouse BAC End Sequences from Library RPCI-23
Unpublished (1999)

TITLE

JOURNAL

COMMENT

Other_GSSs: RPCI-23-15E12.TJ
Contact: Shaying Zhao
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: szhao@tigr.org
Clones are derived from the mouse BAC library RPCI-23. For BAC
library availability, please contact Pieter de Jong
(pieter@edj.med.buffalo.edu). Clones may be purchased from
BACPAC Resources (<http://bacpac.med.buffalo.edu/orderingframe.htm>)
or from Resea ch Genetics (inforesgen.com). BAC end page:
http://www.tigr.org/tdb/bac_ends/mouse/bac_end_intro.html
Plate: 15 row: E column: 12
Seq primer: T7
Class: BAC ends.

FEATURES

source

1. .608

Location/Qualifiers

/organism="Mus musculus"

/strain="C57BL/6J"

/db_xref="taxon:10090"

/clone="RPCI-23-15E12"

/clone_lib="RPCI-23"

/sex="Female"

/lab_host="DH10B"

/note="Organ: Kidney/Brain; Vector: pBACe3.6; Site.1:
EcoRI; Site.2: EcoRI; Female C57BL/6J mouse kidney and/or
brain genomic DNA was isolated and partially digested
with a combination of EcoRI and EcoRI Methylase. Size
selected DNA was cloned into the pBACe3.6 vector at the
EcoRI sites. The ligation products were transformed into
DH10B electrocompetent cells (BRL Life Technologies)."

BASE COUNT

ORIGIN

198 a 97 c 143 g 170 t

Query Match

Best Local Similarity

15.2%; Score 20; DB 17; Length 608;

100.0%; Pred. No. 12;

```

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 94 GAGACCACAGCAGGAGAA 113
|||||
Db 201 GAGACCACAGCAGGAGAA 220

RESULT 5
AG012646
LOCUS Homo sapiens genomic DNA, 21q region, clone: f50C10X10, genomic
survey sequence.
ACCESSION AG012646
VERSION AG012646.1 GI:3419975
KEYWORDS GSS.
SOURCE Homo sapiens DNA, clone:f50C10X10.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 732)
AUTHORS Hattori,M., Ishii,K., Toyoda,A., Shiba,T. and Sakaki,Y.
TITLE Homo sapiens genomic DNA, chromosome 21q
JOURNAL Published Only in Database (1998)
REFERENCE 2 (bases 1 to 732)
AUTHORS Hattori,M., Ishii,K., Toyoda,A., Shiba,T. and Sakaki,Y.
TITLE Direct Submission
JOURNAL Submitted (15-AUG-1998) Masahira Hattori, Kitasato University,
Department of Science, JST Sequencing Laboratory, Kitasato 1-15-1,
Sagamihara 228, Japan (E-mail:hattori@jst.ac.jp,
Tel:0427-78-9732, Fax:0427-78-9561)

FEATURES
source
1..732
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="21"
/map="21q"
/clone="f50C10X10"

BASE COUNT 141 a 162 c 200 g 224 t 5 others
ORIGIN

Query Match 15.2%; Score 20; DB 17; Length 732;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 27 AATGCCAGCTTGGATAAGG 46
|||||
Db 71 AATGCCAGCTTGGATAAGG 90

RESULT 6
BI524256
LOCUS Homo sapiens 997 bp mRNA linear EST 29-AUG-2001
DEFINITION 603052058R1 NIH_MGC_122 Homo sapiens cDNA clone IMAGE:5201607 3',
mRNA sequence.
ACCESSION BI524256
VERSION BI524256.1 GI:15349048
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 997)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cga@r-re@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:

```

```

http://image.llnl.gov
Plate: LLAM11505 row: c column: 16
High quality sequence stop: 68.

FEATURES
source
1..997
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5201607"
/clone_lib="NIH_MGC_122"
/lab_host="DH10B"
/note="organ: pooled lung and spleen; Vector: pCMV-SPORT6;
Site_1: NotI; Site_2: EcoRV (destroyed); RNA source
anonymous pool of 24 week female lung, 16 week female
spleen, and 20-22 week male spleens. Library is oligo-dT
primed and directionally cloned (EcoRV site is destroyed
upon cloning). Average insert size 1.4 kb, insert size
range 1-3 kb. Library is normalized and enriched for
full-length clones and was constructed by C. Gruber
(invitrogen). Research Genetics tracking code 026. Note:
this is a NIH_MGC library."

BASE COUNT 381 a 213 c 207 g 196 t
ORIGIN

Query Match 15.2%; Score 20; DB 13; Length 997;
Best Local Similarity 100.0%; Pred. No. 15;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 93 AGAGACCACAGCAGGAGAA 112
|||||
Db 629 AGAGACCACAGCAGGAGAA 648

RESULT 7
N88692
LOCUS K5134F Human fetal heart, Lambda ZAP Express Homo sapiens cDNA
DEFINITION clone K5134 5' similar to THYMOSIN BETA-10, mRNA sequence.
ACCESSION N88692
VERSION N88692.1 GI:1442022
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 198)
AUTHORS Liew,C.C.
TITLE cDNAs from fetal heart (1996)
JOURNAL Unpublished (1996)
COMMENT Contact: Liew CC
Brigham and Women's Hospital
Harvard Medical School
75 Francis St. Boston, MA 02115, USA
Tel: 617/7328915
Fax: 617/9750995
Email: cliew@rics.bwh.harvard.edu
Seq primer: GAAATTAACCTCACTAAAGG.

FEATURES
source
1..198
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="K5134"
/clone_lib="Human fetal heart, Lambda ZAP Express"
/lab_host="E. coli XL1-Blue"
/note="Vector: Lambda ZAP Express; Site_1: EcoRI; Site_2:
XhoI; mRNA was purified from human fetal hearts (8-10
weeks). cDNA was synthesized using a XhoI-Oligo dT
adaptor-primer. EcoRI adaptors were ligated, followed by
digestion with XhoI, for directional cloning into
predigested lambda ZAP Express."

BASE COUNT 63 a 44 c 60 g 31 t
ORIGIN

Query Match 14.4%; Score 19; DB 14; Length 198;

```

```

Best Local Similarity 100.0%; Pred. No. 27;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 40 GATAAGGCCAAGCTGAAGG 58
Db 87 GATAAGGCCAAGCTGAAGG 105

RESULT 8
BF800160/c
LOCUS
DEFINITION PM0-C10046-181000-001-H07 C10046 Homo sapiens cDNA, mRNA sequence.
ACCESSION BF800160
VERSION BF800160.1 GI:12129149
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 203)
Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.
Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
20202663
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=PM0&st2=PM0-C10046-
181000-001-H07&ts=2000-10-18&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 7
High quality sequence stop: 203.
Location/Qualifiers
1..203
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="C10046"
/dev_stage="Adult"
/note="Organ: colon_ins; Vector: puc18; Site_1: SmaI;
Site_2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."
BASE COUNT 39 a 57 c 68 g 39 t
ORIGIN
Query Match 14.4%; Score 19; DB 12; Length 203;
Best Local Similarity 100.0%; Pred. No. 27;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 92 AAGAGACCACAGCAGGAGGA 110
Db 66 AAGAGACCACAGCAGGAGGA 48

RESULT 9
Az808724
LOCUS

```

```

DEFINITION 2M0072G24F Mouse 10kb plasmid UUGCLM library Mus musculus genomic
clone UUGC2M0072G24 F, DNA sequence.
ACCESSION AZ808724
VERSION AZ808724.1 GI:12974345
KEYWORDS GSS.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 354)
Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
and Wright,D., Weiss,R.
Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts
Unpublished (2000)
Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0072 row: G column: 24
Seq primer: CGTTGTAACGACGCGCCAGT
Class: plasmid ends
High quality sequence stop: 354.
Location/Qualifiers
1..354
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGC2M0072G24"
/clone_lib="Mouse 10kb plasmid UUGCLM library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, TI-resistant, F-"
/note="Vector: PWD42nv; Purified genomic DNA from M.
musculus C57BL/6J (male) was obtained from the Jackson
Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA
was hydrodynamically sheared by repeated passage through a
0.005 inch orifice at constant velocity. The sheared DNA
was blunt end-repaired with T4 DNA polymerase and T4
polynucleotide kinase. Adaptor oligonucleotides were
ligated to the blunt ends in high molar excess. The
adaptored DNA was purified and size-selected for a 9.5 to
10.5 kb range using preparative agarose gel
electrophoresis. Vector DNA was prepared from a derivative
of PWD42 (gil4732114|gb|AF129072.1), a copy-number
inducible derivative of plasmid R1. The vector was ligated
with adaptors complementary to the insert adaptors and
purified. The sheared, adaptored mouse DNA was annealed to
adaptored vector DNA, and transformed into
chemically-competent E. coli XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."
BASE COUNT 104 a 71 c 62 g 117 t
ORIGIN
Query Match 14.4%; Score 19; DB 17; Length 354;
Best Local Similarity 100.0%; Pred. No. 33;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 43 AAGGCCAAGCTGAAGGCCA 61
Db 113 AAGGCCAAGCTGAAGGCCA 131

RESULT 10
Az250549/c
LOCUS

```



```

DEFINITION  RPCI-23-18017.TJB RPCI-23 Mus musculus genomic clone RPCI-23-18017,
DNA sequence.
ACCESSION   A2250549
VERSION     A2250549.1 GI:8563752
KEYWORDS    GSS.
SOURCE      house mouse.
ORGANISM    Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1 (bases 1 to 509)
AUTHORS     Zhao,S., Nierman,W., Feldblyum,T., Malek,J., Shatsman,S., Akinret,
            B., Levins,M., McGann,S., Tsegaye,G., Geer,K., Krol,M., de Jong,P.
            and Fraser,C.M.
TITLE       Mouse BAC End Sequences from Library RPCI-23
JOURNAL     Unpublished (1999)
COMMENT     Other.GSSs: RPCI-23-18017.TV
            Contact: Shaying Zhao
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: szhao@tigr.org
            Clones are derived from the mouse BAC library RPCI-23. For BAC
            library availability, please contact Pieter de Jong
            (pieter@dejong.med.buffalo.edu). Clones may be purchased from
            BACPAC Resources (http://bacpac.med.buffalo.edu/orderingframe.htm)
            or from Resea ch Genetics (info@resgen.com). BAC end page:
            http://www.tigr.org/tdb/bac\_ends/mouse/bac\_end\_intro.html
            Plate: 18 row: 0 Column: 17
            Seq primer: SP6
            Class: BAC ends.

FEATURES             source
    Location/Qualifiers
        1..509
            /organism="Mus musculus"
            /strain="C57BL/6J"
            /db_xref="taxon:10090"
            /clone="RPCI-23-18017"
            /clone_lib="RPCI-23"
            /sex="Female"
            /lab_host="DH10B"
            /note="Organ: Kidney/Brain; Vector: pBACE3.6; Site:1:
            EcoRI; Site_2: EcoRI; Female C57BL/6J mouse kidney and/or
            brain genomic DNA was isolated and partially digested
            with a combination of EcoRI and EcoRI Methylase. Size
            selected DNA was cloned into the pBACE3.6 vector at the
            EcoRI sites. The ligation products were transformed into
            DH10B electrocompetent cells (BRL Life Technologies)."
BASE COUNT  164 a 92 c 110 g 143 t
ORIGIN
Query Match 14.4%; Score 19; DB 17; Length 509;
Best Local Similarity 100.0%; Pred. No. 38;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 43 AAGGCCAAGCTGAAGGCCA 61
|||||
Db 202 AAGGCCAAGCTGAAGGCCA 184

RESULT 11
BJ084935/c
LOCUS
DEFINITION  BJ084935 NIBB Mochii normalized xenopus tailbud library Xenopus
            laevis cDNA clone XL102a03 3', mRNA sequence.
ACCESSION   BJ084935
VERSION     BJ084935.1 GI:17580703
KEYWORDS    EST.
SOURCE      African clawed frog.
ORGANISM    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
            Xenopodinae; Xenopus.

REFERENCE   1 (bases 1 to 530)
AUTHORS     Kitayama,A., Terasaka,C., Mochii,M., Ueno,N., Shin-i,T. and Kohara
            Y.
TITLE       Expressed genes in X. laevis embryo
JOURNAL     Unpublished (2001)
COMMENT     Contact: Tadasu Shin-i
            Center For Genetic Resource Information
            National Institute of Genetics
            1111 Yata, Mishima, Shizuoka 411-8540, Japan
            Tel: 81-559-81-6856
            Fax: 81-559-81-6855
            Email: tshini@genes.nig.ac.jp.
            Location/Qualifiers
                1..530
                    /organism="Xenopus laevis"
                    /db_xref="taxon:8355"
                    /clone="XL102a03"
                    /clone_lib="NIBB Mochii normalized Xenopus tailbud
                    library"
                    /tissue_type="whole embryo"
                    /dev_stage="stage 25"
                    /note="vector: pBSRM3; Site_1: NotI; Site_2: EcoRI; cDNAs
                    were oligo-dt primed and directionally cloned. Staging
                    according to Nieuwkoop and Faber. Library is subtracted
                    and was constructed by N. Garrett and A.M. Zorn,
                    (Wellcome/CRC Institute)."
BASE COUNT  137 a 101 c 98 g 194 t
ORIGIN
Query Match 14.4%; Score 19; DB 13; Length 530;
Best Local Similarity 100.0%; Pred. No. 39;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 39 GGATAGGCCCAAGCTGAAG 57
|||||
Db 530 GGATAGGCCCAAGCTGAAG 512

RESULT 12
AZ769784/c
LOCUS
DEFINITION  1M0570M13R Mouse 10kb plasmid UUGC1M library Mus musculus genomic
            clone UUGC1M0570M13 R, DNA sequence.
ACCESSION   AZ769784
VERSION     AZ769784.1 GI:12890286
KEYWORDS    GSS.
SOURCE      house mouse.
ORGANISM    Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
            1 (bases 1 to 629)
            Contact: Robert B. Weiss
            University of Utah Genome Center
            University of Utah
            Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
            84112, USA
            Tel: 801 585 5606
            Fax: 801 585 7177
            Email: ddunn@genetics.utah.edu
            Insert Length: 10000 Std Error: 0.00
            Plate: 0570 row: M column: 13
            Seq primer: CACACAGGAACAGCTATGACC
            Class: plasmid ends
            High quality sequence stop: 629.
            Location/Qualifiers
                1..629

```


/organism="Mus musculus"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="UUGC1M0570M13"
 /clone_lib="Mouse 10kb plasmid UUGC1M library"
 /sex="Male"
 /lab_host="E. coli strain XL10-Gold, T1-resistant, F-"
 /note="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pWD42 (gll4732114[gb]/AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT 186 a 131 c 134 g 178 t

Query Match 14.4%; Score 19; DB 17; Length 629;
 Best Local Similarity 100.0%; Pred. NO. 41;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 43 AAGGCCAAGCTGAAGGCCA 61
 |||||
 Db 582 AAGGCCAAGCTGAAGGCCA 564

RESULT 13
 BQ803203
 LOCUS
 DEFINITION WHE2834_H05_O10ZS Triticum monococcum vernalized apex cDNA library
 Triticum monococcum cDNA clone WHE2834_H05_O10, mRNA sequence.
 BQ803203
 ACCESSION BQ803203.1 GI:22018172
 VERSION
 KEYWORDS EST.
 SOURCE Triticum monococcum.
 ORGANISM Triticum monococcum

REFERENCE
 AUTHORS Anderson, O.D., Chao, S., Crossman, C., Dubcovsky, J., Echenique, V., Lazo, G.R., Pham, J., Rausch, C.J., Stamova, B., Wilson, C. and Woo, J.
 TITLE The structure and function of the expressed portion of the wheat genomes - Vernalized apex cDNA library from Triticum monococcum Unpublished (2002)
 JOURNAL
 COMMENT Contact: Olin Anderson
 US Department of Agriculture, Agriculture Research Service, Pacific West Area, Western Regional Research Center
 800 Buchanan Street, Albany, CA 94710, USA
 Tel: 5105959773
 Fax: 5105959818
 Email: oanderson@pw.usda.gov

Sequences have been trimmed to remove vector sequence and low quality sequence with phred score less than 20
 Seq primer: SK primer.
 Location/Qualifiers
 1. .655

FEATURES
 source
 /organism="Triticum monococcum"
 /cultivar="G3116"
 /db_xref="taxon:4568"
 /clone="WHE2834_H05_O10"

/clone_lib="Triticum monococcum vernalized apex cDNA library"
 /tissue_type="Vernalized apex"
 /dev_stage="One month old plants"
 /lab_host="E. coli XL0LR"
 /note="Vector: Lambda Uni-ZAP XR, excised phagemid; Site_1: EcoRI; Site_2: XhoI; One-month old plants were subjected to vernalization treatment by placing them in the cold room at 6 C, under 15hr light/9hr dark condition. Total RNA was prepared from apex tissue extracted from plants with no cold treatment; and from plants with 2-week, 4-week and 6-week cold treatment separately. Equal amount of total RNA was pooled from all four samples, a cDNA library was made using pooled polyA RNA and CDNA clones were in vivo excised at the University of California, Davis (V. Echenique, B. Stamova, J. Dubcovsky). Plasmid DNA preparations and DNA sequencing were performed in the OD Anderson lab (all other authors)."

BASE COUNT 159 a 154 c 178 g 164 t

Query Match 14.4%; Score 19; DB 14; Length 655;
 Best Local Similarity 100.0%; Pred. NO. 42;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 82 CTGATGACCAAGAGACCA 100
 |||||
 Db 535 CTGATGACCAAGAGACCA 553

RESULT 14
 BF317304
 LOCUS
 DEFINITION 601904069F1 NIH_MGC_19 Homo sapiens cDNA clone IMAGE:4136561 5', mRNA
 mRNA sequence.
 BF317304
 ACCESSION BF317304.1 GI:11265629
 VERSION
 KEYWORDS EST.
 SOURCE human.

ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE
 AUTHORS NIH-MGC http://mgc.nci.nih.gov/
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished (1999)
 JOURNAL
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapps-femail.nih.gov

Tissue Procurement: ATCC
 CDNA Library Preparation: Ling Hong/Rubin Laboratory
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
 Plate: L1CM1042 row: j column: 18
 High quality sequence stop: 651.
 Location/Qualifiers
 1. .678

FEATURES
 source

/organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:4136561"
 /clone_lib="NIH_MGC_19"
 /tissue_type="neuroblastoma"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: brain; Vector: pOTB7; Site_1: XhoI; Site_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."

BASE COUNT 120 a 178 c 255 g 125 t

ORIGIN

Query Match 14.4%; Score 19; DB 12; Length 676;
Best Local Similarity 100.0%; Pred. No. 42;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 94 GAGACCACAGACAGAGAGA 112
|||||
Db 611 GAGACCACAGACAGAGAGA 629

RESULT 15
BE882559/c

LOCUS BE882559 724 bp mRNA linear EST 20-OCT-2000
DEFINITION G01507007F1 NIH_MGC_71 Homo sapiens cDNA clone IMAGE:3908904 5',
mRNA sequence.

ACCESSION BE882559
VERSION BE882559.1 GI:10331335
KEYWORDS EST.
SOURCE human.

ORGANISM

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

AUTHORS 1 (bases 1 to 724)
TITLE NIH-MGC http://mgc.nci.nih.gov/.
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)

Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov

Tissue Procurement: ATCC

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

Clone Sequencing by: Incyte Genomics, Inc.

Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: LIAM9721 row: m column: 01

High quality sequence stop: 716.

Location/Qualifiers

FEATURES

source

1..724
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3908904"
/clone_lib="NIH_MGC_71"
/tissue_type="leiomyosarcoma"
/lab_host="DH10B (phage-resistant)"
/note="Organ: uterus; Vector: pCMV-SPORT6; Site.1: NotI;
Site.2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 2.1 kb."
176 a 180 c 228 g 140 t

BASE COUNT

Query Match 14.4%; Score 19; DB 12; Length 724;
Best Local Similarity 100.0%; Pred. No. 43;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 50 AGCTGAAGGCCACAGAGAT 68
|||||
Db 606 AGCTGAAGGCCACAGAGAT 588

RESULT 16
CNS052GR/c

LOCUS CNS052GR 1058 bp DNA linear GSS 26-JUL-2000
DEFINITION Tetraodon nigroviridis genome survey sequence SP6 end of clone
036F09 of library B from Tetraodon nigroviridis, genomic survey
sequence.

AL318132

AL318132.1 GI:9551016

GSS: genome survey sequence.

Tetraodon nigroviridis.

Tetraodon nigroviridis

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

1058 bp DNA linear GSS 26-JUL-2000

Tetraodon nigroviridis genome survey sequence SP6 end of clone

036F09 of library B from Tetraodon nigroviridis, genomic survey

sequence.

AL318132

AL318132.1 GI:9551016

GSS: genome survey sequence.

Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
Tetraodontidae; Tetraodon.

1 (bases 1 to 1058)

Roest Crolius,H., Jaillon,O., Dasilva,C., Bouneau,L., Fisher,C.,

Bernot,A., Fizames,C., Wincker,P., Brottier,P., Quetier,F.,

Saurin,W. and Weissenbach,J.

Estimate of human gene number provided by genome-wide analysis

using Tetraodon nigroviridis DNA sequence

Nat. Genet. 25 (2), 235-238 (2000)

20296633

10835645

2 (bases 1 to 1058)

Crolius,H.R., Jaillon,L., Billault,A., Quetier,F., Saurin,W.,

Fischer,C., Bouneau,L., Billault,A., Quetier,F., Saurin,W.,

Bernot,A. and Weissenbach,J.

Characterization and repeat analysis of the compact genome of the

freshwater pufferfish Tetraodon nigroviridis

Genome Res. 10 (7), 939-949 (2000)

20359837

3 (bases 1 to 1058)

Genoscope.

Submitted (12-APR-2000)

Direct Submission

This sequence is a single read and was generated as part of a large

scale clone-end sequencing project of the Tetraodon nigroviridis

genome. For more information, please take a look at

http://www.genoscope.cns.fr/Tetraodon.

Location/Qualifiers

1..1058

/organism="Tetraodon nigroviridis"

/db_xref="taxon:99883"

/clone="036F09"

/clone_lib="B"

/note="Genoscope sequence ID : C0AB036C005B1-end : SP6"

329 a 233 c 200 g 265 t 31 others

BASE COUNT

Query Match 14.4%; Score 19; DB 17; Length 1058;

Best Local Similarity 100.0%; Pred. No. 50;

Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 22 GAAGAAATGCCAGCTTGG 40

|||||

Db 315 GAAGAAATGCCAGCTTGG 297

RESULT 17

AA096439

19794.seq.F Human fetal heart, Lambda ZAP Express Homo sapiens cDNA

5', mRNA sequence.

AA096439

AA096439.1 GI:1642024

EST.

human.

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 135)

Liew/C.C.

cDNAs from fetal heart (1996)

Unpublished (1996)

Contact: Liew CC

Brigham and Women's Hospital

Harvard Medical School

75 Francis St. Boston, MA 02115, USA

Tel: 6177328915

Fax: 6179750995

Email: cliw@rics.bwh.harvard.edu

PCR Primers

FORWARD: 5' GCCAAGCTCGAANTTACCTCTACTAAAGG 3'

BACKWARD: 5' CCAAGTAATGTAAACGACCTCACTATAGGGCG 3'
Seq primer: 5' GAATTAACCTCACTAAAGG 3'.

FEATURES

Location/Qualifiers
1..135

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="Human fetal heart, Lambda ZAP Express"
/lab_host="E. coli XL1-Blue"
/note="Vector: Lambda ZAP Express; Site_1: EcoRI; Site_2: XhoI; mRNA was purified from human fetal hearts (8-10 weeks). cDNA was synthesized using a XhoI-Oligo dT adaptor-primer. EcoRI adaptors were ligated, followed by digestion with XhoI, for directional cloning into predigested lambda ZAP Express."

BASE COUNT 42 a 30 c 44 g 19 t

ORIGIN

Query Match 13.6%; Score 18; DB 9; Length 135;
Best Local Similarity 100.0%; Pred. No. 77;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 40 GATAGGCCCAAGCTGAAG 57

DB 99 GATAGGCCCAAGCTGAAG 116

RESULT 18

AA623585

LOCUS

DEFINITION vn73d02.r1 Knowles Solter mouse blastocyst B1 Mus musculus CDNA clone IMAGE:1037571 5' similar to gb:S54005 THYMOSIN BETA-10 (HUMAN); mRNA sequence.

ACCESSION

VERSION AA623585

KEYWORDS

EST. AA623585.1 GI:2527461

SOURCE

house mouse.

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 144)
Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T., Geisel, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M., Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B., Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and Waterston, R.

The WashU-HMI Mouse EST Project

Unpublished (1996)

Contact: Marra M/Mouse EST Project

WashU-HMI Mouse EST Project

Washington University School of MedicineP

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

Tel: 314 286 1800

Fax: 314 286 1810

Email: mouseest@watson.wustl.edu

This clone is available royalty-free through LLNL; contact the

IMAGE Consortium (info@image.llnl.gov) for further information.

MGI:577595

Trace considered overall poor quality

High quality sequence stop: 1.

Location/Qualifiers

1..144

/organism="Mus musculus"

/strain="C57BL/6J x DBA/2J F1"

/db_xref="taxon:10090"

/clone="IMAGE:1037571"

/clone_lib="Knowles Solter mouse blastocyst B1"

/tissue_type="blastocyst"

/dev_stage="embryo (pre-implantation)"

/lab_host="DH10B"

/note="Organ: embryo; Vector: pSPORT; Site_1: NotI;

Site_2: SalI; Cloned unidirectionally from mRNA prepared

from 800 blastocysts. Primer: SalI(dT):

5'-CGGTCAGCCGTCACCGCTTTT-3'. cDNAs were

cloned into the NotI/SalI sites of a pSPORT vector (Life Technologies). Two different size selections: B1 (larger inserts) and B3.

BASE COUNT 56 a 29 c 40 g 19 t

ORIGIN

Query Match 13.6%; Score 18; DB 9; Length 144;

Best Local Similarity 100.0%; Pred. No. 79;

Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 40 GATAGGCCCAAGCTGAAG 57

DB 49 GATAGGCCCAAGCTGAAG 66

RESULT 19

AL156792

LOCUS

DEFINITION ue54f03.r1 Soares_mammary_gland_NLMG Mus musculus cDNA clone IMAGE:1494941 5' similar to gb:S54005 THYMOSIN BETA-10 (HUMAN); mRNA sequence.

ACCESSION

VERSION AL156792

KEYWORDS

EST. AL156792.1 GI:3685261

SOURCE

house mouse.

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 154)
Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T., Geisel, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M., Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B., Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and Waterston, R.

The WashU-HMI Mouse EST Project

Unpublished (1996)

Contact: Marra M/Mouse EST Project

WashU-HMI Mouse EST Project

Washington University School of MedicineP

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

Tel: 314 286 1800

Fax: 314 286 1810

Email: mouseest@watson.wustl.edu

This clone is available royalty-free through LLNL; contact the

IMAGE Consortium (info@image.llnl.gov) for further information.

MGI:932545

Trace considered overall poor quality

Seq primer: -28ml3 rev2 ET from Amersham

High quality sequence stop: 1.

Location/Qualifiers

1..154

/organism="Mus musculus"

/db_xref="taxon:10090"

/clone="IMAGE:1494941"

/clone_lib="Soares_mammary_gland_NLMG"

/sex="female (lactating)"

/tissue_type="mammary gland"

/lab_host="DH10B"

/note="vector: pT73p-Pac (Pharmacia) with a modified

polylinker; 1st strand cDNA was prepared from mammary

gland tissue from a lactating female, and was then primed

with a Not I - oligo(dT) primer. Double-stranded cDNA was

ligated to Eco RI adaptors (Pharmacia), digested with Not

I and cloned into the Not I and Eco RI sites of the

modified pT73 vector. Library is normalized. Library

was constructed by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 60 a 33 c 42 g 19 t

ORIGIN

Query Match 13.6%; Score 18; DB 9; Length 154;

Best Local Similarity 100.0%; Pred. No. 81;

Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 40 GATAAGGCCAAGCTGAAG 57
 D76516
 Db 43 GATAAGGCCAAGCTGAAG 60

RESULT 20

AA967821 157 bp mRNA linear EST 19-MAY-1998
 LOCUS un05f03.r1 Soares mouse hypothalamus NMHY Mus musculus cDNA clone
 DEFINITION IMAGE:1617053 5' similar to gb:S54005 THYMOSIN BETA-10 (HUMAN);,
 mRNA sequence.

ACCESSION

AA967821
 VERSION AA967821.1 GI:3141714

KEYWORDS

EST.

SOURCE

house mouse.

ORGANISM

Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1. (Bases 1 to 157)
 Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
 Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
 Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
 Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
 Waterston,R.

TITLE

The WashU-HHMI Mouse EST Project

JOURNAL

Unpublished (1996)

COMMENT

Contact: Marra M/Mouse EST Project
 WashU-HHMI Mouse EST Project
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810

Email: mouseest@watson.wustl.edu

This clone is available royalty-free through LLNL ; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.

MGI:956353

Trace considered overall poor quality

Seq primer: -28ml3 rev2 Et from Amersham

High quality sequence stop: 1.

FEATURES

Location/Qualifiers

1..157

/organism="Mus musculus"

/db_xref="taxon:10090"

/clone="IMAGE:1617053"

/clone_lib="Soares mouse hypothalamus NMHY"

/tissue_type="hypothalamus"

/lab_host="DH10B"

/note="Organ: brain; Vector: pT7T3D-Pac (Pharmacia) with a
 modified polylinker; Site.1: Not I; Site.2: Eco RI; 1st
 strand cDNA was primed with a Not I - oligo(dT) primer [5'
 TGTTACCACTGAAGTGGAGCGCGCAAGGTTTTTTTTTTTTTTTTTTT
 T 3']; double-stranded cDNA was ligated to Eco RI adaptors
 (Pharmacia), digested with Not I and cloned into the Not I
 and Eco RI sites of the modified pT7T3 vector. RNA
 provided by Dr. Wolfgang Liedtke. Library went through
 two rounds of normalization, and was constructed by Bento
 Soares and M.Fatima Bonaldo."

63 a 33 c 43 g 18 t

BASE COUNT

ORIGIN

Query Match

Best Local Similarity 13.6%; Score 18; DB 9; Length 157;

Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY

40 GATAAGGCCAAGCTGAAG 57

Db

46 GATAAGGCCAAGCTGAAG 63

RESULT 21

D76516 157 bp mRNA linear EST 07-OCT-1996

LOCUS

DEFINITION MUS70C08 mouse embryonal carcinoma cell line F9 Mus musculus cDNA

ACCESSION
 VERSION
 KEYWORDS
 SOURCE

ORGANISM

clone 70C08, mRNA sequence.
 D76516
 D76516.1 GI:1596294

EST.

house mouse.

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE

1 (Bases 1 to 157)

Nishiguchi,S., Sakuma,R., Nomura,M., Zou,Z., Jearanaisilavong,J.,
 Joh,T., Yasunaga,T. and Shimada,K.

A catalogue of genes in mouse embryonal carcinoma F9 cells
 identified with expressed sequence tags

J. Biochem. 119 (4), 749-767 (1996)

96337530

COMMENT

Contact: Kazunori Shimada

Department of Medical Genetics, Division of Molecular Biomedicine

Research Institute for Microbial Diseases, Osaka University

3-1, Yamadaoka, Suita, Osaka, 565, Japan

Tel: 06-879-8325

Fax: 06-879-8326.

FEATURES

Location/Qualifiers

1..157

/organism="Mus musculus"

/db_xref="taxon:10090"

/clone="70C08"

/clone_lib="mouse embryonal carcinoma cell line F9"

/note="Vector: Uni-Zap XR; Site.1: EcoR I; Site.2: Xho I;
 mRNA was purified from 14.7 hybrid MN cells and used to
 construct a size- selected unidirectional cDNA library in
 Uni-Zap XR vector."

55 a 38 c 46 g 15 t 3 others

BASE COUNT

ORIGIN

Query Match

Best Local Similarity 13.6%; Score 18; DB 14; Length 157;

Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY

40 GATAAGGCCAAGCTGAAG 57

Db

90 GATAAGGCCAAGCTGAAG 107

RESULT 22

AA792584

LOCUS

DEFINITION

AA792584 158 bp mRNA linear EST 09-FEB-1998

VS88e04.r1 Barstead mouse myotubes MPLRB5 Mus musculus cDNA clone

IMAGE:1153374 5' similar to gb:S54005 THYMOSIN BETA-10 (HUMAN);,
 mRNA sequence.

AA792584

VERSION

AA792584.1 GI:2855539

KEYWORDS

EST.

SOURCE

ORGANISM

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE

1 (Bases 1 to 158)

Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
 Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
 Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
 Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
 Waterston,R.

The WashU-HHMI Mouse EST Project

Unpublished (1996)

Contact: Marra M/Mouse EST Project

WashU-HHMI Mouse EST Project

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

Tel: 314 286 1800

Fax: 314 286 1810

EMAIL: mouseest@watson.wustl.edu

This clone is available royalty-free through LLNL ; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.


```

/tissue_type="fetus"
/dev_stage="12.5dpc total fetus"
/lab_host="DH10B"
/Note="Organ: whole fetus; Vector: pT7T3D-Pac (Pharmacia)
with a modified polylinker; Site_1: Not I; Site_2: Eco RI;
1st strand cDNA was primed with a Not I - oligo(dT) primer
[5' TGTTACCAATCTGAAGTGGAGCGGCCCTATTTTTTTTTTTTTTTT
3'], on total mouse RNA [provided by Minoru Ko, Wayne
State Univ.]; double-stranded cDNA was ligated to Eco RI
adaptors (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of the modified pT7T3 vector.
Library went through one round of normalization, and was
constructed by Bento Soares and M. Fatima Bonaldo. "
```

59 a 40 c 43 g 18 t

BASE COUNT 59 a 40 c 43 g 18 t

ORIGIN

Query Match 13.6%; Score 18; DB 9; Length 160;

Best Local Similarity 100.0%; Pred. No. 82;

Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 40 GATAAGGCCAAGCTGAAG 57

Db 40 GATAAGGCCAAGCTGAAG 57

RESULT 25

AA207933

LOCUS

DEFINITION

mv92c01.r1 GuayWoodford Beier mouse kidney day 7 Mus musculus cDNA

clone IMAGE:662496 5' similar to gb:S54005 THYMOSIN BETA-10 (HUMAN

);, mRNA sequence.

AA207933 161 bp mRNA linear EST 12-MAR-1997

VERSION mv92c01.r1 GuayWoodford Beier mouse kidney day 7 Mus musculus cDNA

KEYWORDS

SOURCE

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,

Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,

Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,

Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and

Waterston,R.

The WashU-HHMI Mouse EST Project

Unpublished (1996)

Contact: Marra M/Mouse EST Project

WashU-HHMI Mouse EST Project

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

Tel: 314 286 1800

Fax: 314 286 1810

Email: mouseest@watson.wustl.edu

This clone is available royalty-free through LLNL; contact the

IMAGE Consortium (info@image.llnl.gov) for further information.

MGT:408344

Seq primer: -28ml3 rev1 ET from Amersham

High quality sequence stop: 77.

Location/Qualifiers

1. 161

/organism="Mus musculus"

/strain="C57BL/6J"

/db_xref="taxon:10090"

/clone="IMAGE:662496"

/clone_lib="GuayWoodford Beier mouse kidney day 7"

/tissue_type="kidney"

/dev_stage="juvenile (7 days old)"

/lab_host="SOLR (kanamycin resistant)"

/Note="Organ: kidney; Vector: pBluescript SK-; Site_1:

EcoRI; Site_2: XhoI; Cloned unidirectionally. Primer:

Oligo dT. Average insert size: 1.0 kb; Uni-ZAP XR Vector;

-5' adaptor sequence: 5' GAATTCGCGACGAG 3' -3' adaptor

sequence: 5' CTCGAGTTTTTTTTTTTTTTTTTTT 3' Library provided

```

by Lisa Guay Woodford. "
62 a 33 c 45 g 21 t


BASE COUNT 62 a 33 c 45 g 21 t



ORIGIN



Query Match 13.6%; Score 18; DB 9; Length 161;



Best Local Similarity 100.0%; Pred. No. 82;



Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;



QY 40 GATAAGGCCAAGCTGAAG 57



Db 49 GATAAGGCCAAGCTGAAG 66



RESULT 26



AA792947



LOCUS



DEFINITION



vp28e01.r1 Barstead mouse proximal colon MPLRB6 Mus musculus cDNA



clone IMAGE:1078008 5' similar to gb:S54005 THYMOSIN BETA-10 (HUMAN



);, mRNA sequence.



AA792947 162 bp mRNA linear EST 09-FEB-1998



VERSION vp28e01.r1 Barstead mouse proximal colon MPLRB6 Mus musculus cDNA



KEYWORDS



SOURCE



ORGANISM



Mus musculus



house mouse.



Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;



Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.



Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,



Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,



Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,



Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and



Waterston,R.



The WashU-HHMI Mouse EST Project



Unpublished (1996)



Contact: Marra M/Mouse EST Project



WashU-HHMI Mouse EST Project



Washington University School of Medicine



4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108



Tel: 314 286 1800



Fax: 314 286 1810



Email: mouseest@watson.wustl.edu



This clone is available royalty-free through LLNL; contact the



IMAGE Consortium (info@image.llnl.gov) for further information.



MGT:592304



Seq primer: -28ml3 rev2 ET from Amersham



High quality sequence stop: 107.



Location/Qualifiers



1. 162



/organism="Mus musculus"



/strain="FVB/N"



/db_xref="taxon:10090"



/clone="IMAGE:1078008"



/clone_lib="Barstead mouse proximal colon MPLRB6"



/dev_stage="7 day juvenile"



/lab_host="DH10B"



/Note="Vector: pT7T3D-Pac (Pharmacia) with a modified



polylinker; Site_1: EcoRI; Site_2: NotI; 1st strand cDNA



was primed with a Not I - oligo(dT) primer [5'



TGTTACCAATCTGAAGTGGAGCGGCCCTATTTTTTTTTTTTTTTT



3']; double-stranded cDNA was ligated to Eco RI adaptors



[AATTGCGATCCTTG], digested with Not I and cloned into the



Not I and Eco RI sites of the modified pT7T3 vector.



Library constructed by Bob Barstead.



62 a 35 c 45 g 20 t



BASE COUNT 62 a 35 c 45 g 20 t



ORIGIN



Query Match 13.6%; Score 18; DB 9; Length 162;



Best Local Similarity 100.0%; Pred. No. 82;



Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;



QY 40 GATAAGGCCAAGCTGAAG 57



Db 49 GATAAGGCCAAGCTGAAG 66


```

```

RESULT 27
AA220778
LOCUS
DEFINITION
  AA220778 162 bp mRNA linear EST 10-FEB-1997
  mv9c04.r1 Soares mouse 3NME12.5 Mus musculus cDNA clone
  IMAGE:660294 5' similar to gb:S54005 THYMOSIN BETA-10 (HUMAN);,
  mRNA sequence.
ACCESSION
AA220778
VERSION
AA220778.1 GI:1838598
KEYWORDS
EST.
SOURCE
house mouse.
ORGANISM
Mus musculus
REFERENCE
1 (bases 1 to 162)
AUTHORS
Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
Waterston,R.
TITLE
The WashU-HMI Mouse EST Project
JOURNAL
Unpublished (1996)
COMMENT
Contact: Marra M/Mouse EST Project
WashU-HMI Mouse EST Project
Washington University School of MedicineP
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LLNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:406142
Trace considered overall poor quality
Seq primer: -28ml3 rev2 ET from Amersham
High quality sequence stop: 1.
Location/Qualifiers
1. .162
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="IMAGE:660294"
/clone_lib="Soares mouse 3NME12 5"
/sex="unknown"
/tissue_type="fetus"
/dev_stage="12.5dpc total fetus"
/lab_host="PH108"
/note="Organ: whole fetus; Vector: pT73D-Pac (Pharmacia)
with a modified polylinker; Site_1: Not I; Site_2: Eco RI;
1st strand cDNA was primed with a Not I - oligo(dT) primer
[5' TGTTACCAATCTGAAGTGGGCGCGCTTATTTTATTTTATTTT
3'], on total mouse RNA [provided by Minoru Ko, Wayne
State Univ.]; double-stranded cDNA was ligated to Eco RI
adaptors (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of the modified pT73 vector.
Library went through one round of normalization, and was
constructed by Bento Soares and M. Fatima Bonaldo."
BASE COUNT 63 a 34 c 45 g 20 t
ORIGIN
Query Match 13.6%; Score 18; DB 9; Length 162;
Best Local Similarity 100.0%; Pred. No. 82;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 40 GATAAGGCCAAGCTGAAG 57
|||||
DB 49 GATAAGGCCAAGCTGAAG 66
|||||

RESULT 28
AA591963
LOCUS
DEFINITION
  AA591963 162 bp mRNA linear EST 16-SEP-1997
  vi50b08.r1 Beddington mouse embryonic region Mus musculus cDNA
  IMAGE:907191 5' similar to gb:S54005 THYMOSIN BETA-10 (HUMAN);,
  mRNA sequence.
ACCESSION
AA591963
VERSION
AA591963.1 GI:2405626
KEYWORDS
EST.
SOURCE
house mouse.
ORGANISM
Mus musculus
REFERENCE
1 (bases 1 to 162)
AUTHORS
Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
Waterston,R.
TITLE
The WashU-HMI Mouse EST Project
JOURNAL
Unpublished (1996)
COMMENT
Contact: Marra M/Mouse EST Project
WashU-HMI Mouse EST Project
Washington University School of MedicineP
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LLNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:527855
Trace considered overall poor quality
Seq primer: -28ml3 rev1 ET from Amersham
High quality sequence stop: 1.
Location/Qualifiers
1. .162
/organism="Mus musculus"
/strain="C57BL/6 x DBA"
/db_xref="taxon:10090"
/clone="IMAGE:907191"
/clone_lib="Beddington mouse embryonic region"
/sex="pooled"
/tissue_type="embryo"
/dev_stage="7.5dpc"
/lab_host="DH12S"
/note="Organ: whole embryo; Vector: pCMV-SPORT; Site_1:
SalI; Site_2: NotI; Cloned unidirectionally. Primer:
Oligo dT. Gastrulating embryos were collected at 7.5dpc
from C57BL6 x DBA matings, excluding embryos that had
developed head folds and all extraembryonic tissues.
Average insert size: 1.3 kb (range: 0.5 - 3.0 kb).
Referenced in Development 121, 2479-2489 (1995)"
BASE COUNT 65 a 38 c 44 g 15 t
ORIGIN
Query Match 13.6%; Score 18; DB 9; Length 162;
Best Local Similarity 100.0%; Pred. No. 82;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 40 GATAAGGCCAAGCTGAAG 57
|||||
DB 49 GATAAGGCCAAGCTGAAG 66
|||||

RESULT 29
BM843852
LOCUS
DEFINITION
  BM843852 163 bp mRNA linear EST 06-MAR-2002
  K-EST0121847 SI2SNU216 Homo sapiens cDNA clone SI2SNU216-58-D10 5',
  mRNA sequence.
ACCESSION
BM843852
VERSION
BM843852.1 GI:19200261
KEYWORDS
EST.
SOURCE
human.
ORGANISM
Homo sapiens
REFERENCE
1 (bases 1 to 163)

```

AUTHORS

Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
Kim,Y.S.

TITLE
JOURNAL
COMMENT

21C Frontier Korean EST Project 2001
Unpublished (2002)
Contact: Kim YS

Genome Research Center

Korea Research Institute of Bioscience & Biotechnology
52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea

Tel: +82-42-860-4470

Fax: +82-42-860-4409

Email: yongsung@mail.kribb.re.kr

Plate: 58 row: D column: 10

High quality sequence stop: 163.

FEATURES

source

1. .163
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="S125N0216-58-D10"
/clone_lib="S125N0216"
/sex="F"
/tissue_type="Lymph node"
/cell_type="Epithelial"
/cell_lines="SMU-216"
/lab_host="Top10F"
/note="Organ: Stomach; Vector: pCNS; Site_1: EcoRI;
Site_2: NotI; The poly (A)+ RNA was dephosphorylated with
bacterial alkaline phosphatase (BAP) and then deacapped
with tabacco acid pyrophosphatase (TAP). The deacapped
intact mRNA was ligated with DNA-RNA linker including EcoR
I site by treatment of 14 RNA ligase and the first strand
cDNA was synthesized from oligo dt-selected mRNA by
priming with dt-tailed vector. The dt-tailed vector was
adjusted to have about 60nt. The cDNA vector was
circularized with E. coli DNA ligase after digestion of
EcoRI which site is also included in vector. An RNA strand
converted to a DNA strand by Okayama-Berg method. The
obtained cDNA vectors were used for transformation of
competent cells E. coli Top10F, by electroporation method.
The cDNA libraries constructed by this method are
full-length enriched cDNA library."

53 a 34 c 52 g 24 t
BASE COUNT
ORIGIN

Query Match 13.6%; Score 18; DB 14; Length 163;

Best Local Similarity 100.0%; Pred. No. 83;

Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 40 GATAAGGCCAAGCTGAAG 57

|||||

Db 119 GATAAGGCCAAGCTGAAG 136

RESULT 30

BM686528

LOCUS

UI-E-CQ0-adn-f-08-0-UI.r1 164 bp mRNA linear EST 28-FEB-2002

UI-E-CQ0-adn-f-08-0-UI 5', mRNA sequence.

BM686528

VERSION

BM686528.1 GI:18999786

KEYWORDS

EST.

SOURCE

human.

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 164)

Bonaldo,M.F., Lennon,G. and Soares,M.B.

Normalization and subtraction: two approaches to facilitate gene

discovery

Genome Res. 6 (9), 791-806 (1996)

JOURNAL

MEDLINE

97044477

COMMENT

Contact: Soares, MB

Program for Rat Gene Discovery and Mapping

University of Iowa

451 Eckstein Medical Research Building Iowa City, IA 52242, USA

Tel: 319 335 8250

Fax: 319 335 9565

Email: msoares@blue.weeg.uiowa.edu

Tissue Procurement: Dr. Gregg Hageman

cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa

cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa

DNA Sequencing by: Dr. M. Bento Soares, University of Iowa

Clone Distribution: Researchers may obtain clones from Research

Genetics (www.resgen.com).

Seq primer: M13 Reverse.

Location/Qualifiers

1. .164

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="UI-E-CQ0-adn-f-08-0-UI"

/clone_lib="UI-E-CQ0"

/tissue_type="optic nerve"

/dev_stage="adult"

/lab_host="DH10B (Life Technologies) (T1 phage resistant)"

/note="Organ: eye; Vector: pT73-Pac (Pharmacia) with a

modified polylinker; Site_1: EcoR I; Site_2: Not I;

UI-E-CQ0 is a cDNA library containing the following

tissue(s): optic nerve. The library was constructed

according to Bonaldo, Lennon and Soares, Genome Research,

6:791-806, 1996. First strand cDNA synthesis was primed

with an oligo-dT primer containing a Not I site. Double

stranded cDNA was ligated to an EcoR I adaptor, digested

with Not I, and cloned directionally into pT73-Pac

vector. The oligonucleotide used to prime the synthesis of

first-strand cDNA contains a library tag sequence that is

located between the Not I site and the (dT)18 tail. The

sequence tag for this library is CCATTAAGTG. This library

was created for the program, Gene Discovery in the Visual

System, supported by National Eye Institute (NEI)."

55 a 38 c 52 g 19 t

BASE COUNT

ORIGIN

Query Match 13.6%; Score 18; DB 14; Length 164;

Best Local Similarity 100.0%; Pred. No. 83;

Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 40 GATAAGGCCAAGCTGAAG 57

|||||

Db 107 GATAAGGCCAAGCTGAAG 124

RESULT 31

AA144770

LOCUS

DEFINITION

mr68a05.r1 Stratagene mouse testis (#937308) Mus musculus cDNA

clone IMAGE:602576 5' similar to gb:S54005 THYMOSIN BETA-10 (HUMAN

):: mRNA sequence.

AA144770

VERSION

AA144770.1 GI:1714141

KEYWORDS

EST.

SOURCE

house mouse.

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 169)

Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,

Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,

Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,

Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and

Waterston,R.

The WashU-HMI Mouse EST Project

Unpublished (1996)

Contact: Marra M/Mouse EST Project

WashU-HMI Mouse EST Project

Washington University School of MedicineP

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108


```

Tel: 314 286 1800
Fax: 314 286 1810
Email: mouse@wustl.edu
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.lnl.gov) for further information.
MGI:368008
Seq primer: -28m13 rev1 ET from Amersham
High quality sequence stop: 1.
Location/Qualifiers
1. 169
/organism="Mus musculus"
/strain="CD-1"
/db_xref="taxon:10090"
/clone="IMAGE:602576"
/clone_lib="Stratagene mouse testis (#937308)"
/sex="males"
/tissue="testis"
/dev_stage="10-12 week old"
/lab_host="SOLR (kanamycin resistant)"
/Note="Organ: testis; Vector: pBluescript SK-; Site_1:
EcoRI; Site_2: XhoI; Cloned unidirectionally. Primer:
Oligo dt. Average insert size: 1.0 kb; Uni-ZAP XR Vector;
-5' adaptor sequence: 5' GAATTCGACGAG 3' -3' adaptor
sequence: 5' CTCGAGTTTTTTTTTTTTTTT 3'"
BASE COUNT      65 a 41 c 43 g 20 t
ORIGIN
Query Match      13.6%; Score 18; DB 9; Length 169;
Best Local Similarity 100.0%; Pred. No. 84;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 40 GATAAGGCCAAGCTGAAG 57
|||||
Db 49 GATAAGGCCAAGCTGAAG 66
|||||

RESULT 32
BE774336
LOCUS
DEFINITION
MRI-UM0009-050600-021-a06 UM0009 Homo sapiens cDNA, mRNA sequence.
ACCESSION
BE774336
VERSION
BE774336.1 GI:10227991
KEYWORDS
EST.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 169)
Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
Brunstein,A., deoliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.
Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
20202663
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=et2-mri-UM0009-050
600-021-a06&t3=2000-06-05&t4=1)
Seq primer: puc 18 forward
High quality sequence stop: 169.
Location/Qualifiers

```

```

1. 169
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="UM0009"
/dev_stage="Adult"
/Note="Organ: uterus; Vector: puc18; Site_1: SmaI; Site_2:
SmaI; A mini-library was made by cloning products derived
from ORESTES PCR (U.S. Letters Patent application No. 196
,716 - Ludwig Institute for Cancer Research) profiles
into the pUC 18 vector. Reverse transcription of tissue
mRNA and cDNA amplification were performed under low
stringency conditions."
BASE COUNT      51 a 43 c 48 g 27 t
ORIGIN
Query Match      13.6%; Score 18; DB 12; Length 169;
Best Local Similarity 100.0%; Pred. No. 84;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 40 GATAAGGCCAAGCTGAAG 57
|||||
Db 22 GATAAGGCCAAGCTGAAG 39
|||||

RESULT 33
R57257
LOCUS
DEFINITION
F2310 Fetal heart Homo sapiens cDNA clone F2310 5' end similar to
Thymosin beta-10, mRNA sequence.
ACCESSION
R57257
VERSION
R57257.1 GI:827315
KEYWORDS
EST.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 170)
Hwang,D.M., Fung,Y.W., Wang,R.X., Laurensen,C.M., Ng,S.H., Lam
,W.Y., Tsui,K.W., Fung,K.P., Waye,M., Lee,C.Y. and Liew,C.C.
Analysis of expressed sequence tags from a fetal human heart cDNA
library
Genomics 30 (2), 293-298 (1995)
96163883
Contact: Liew CC
Brigham and Women's Hospital
Harvard Medical School
75 Francis St. Boston, MA 02115, USA
Tel: 6177328915
Fax: 6179750995
Email: cillewerts.bwh.harvard.edu
Seq primer: GTGGCGACGACTCCTGGAGCC.
Location/Qualifiers
1. 170
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="F2310"
/clone_lib="Fetal heart"
/lab_host="E. coli Y1090"
/Note="Vector: Lambda gt22; Site_1: NotI; Site_2: SalI;
mRNA was purified from human fetal hearts (10-12 weeks).
cDNA was constructed using a NotI-oligo dt adaptor-primer.
Sali adaptors were ligated, followed by digestion with
NotI, for direction cloning into predigested lambda gt22.
Method is described in J. Mol. Cell. Cardiol. (1994) 26,
1329-1333,"
BASE COUNT      55 a 44 c 46 g 24 t 1 others
ORIGIN
Query Match      13.6%; Score 18; DB 14; Length 170;
Best Local Similarity 100.0%; Pred. No. 84;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 40 GATAAGGCCAAGCTGAAG 57

```

```

Db
|||||
2 GATAAGGCCAAGCTGAAG 19

RESULT 34
AA375556
LOCUS
DEFINITION
  EST87853 HSC172 cells II Homo sapiens cDNA 5' end similar to
  similar to thymosin beta-10 (GB:S54005), mRNA sequence.
ACCESSION
  AA375556
VERSION
  AA375556.1 GI:2027876
KEYWORDS
  EST.
ORGANISM
  human.
REFERENCE
  Homo sapiens
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
  1 (bases 1 to 171)
  Adams,M.D., Kervage,A.R., Fleischmann,R.D., Fuldner,R.A., Bult
  ,C.J., Lee,N.H., Kirkness,E.F., Weinstock,K.G., Gocayne,J.D., White
  ,O., Sutton,G., Blake,J.A., Brandon,R.C., Man-Wai,C., Clayton,R.A.,
  Cline,T.R., Cotton,M.D., Earle-Hughes,J., Fine,L.D., Fitzgerald
  ,L.M., Fitzhugh,W.M., Fritchman,J.L., Geohagen,N.S., Glodek,A.,
  Gnehm,C.L., Hanna,M.C., Hedblom,E., Hinkle,P.S.Jr., Kelley,J.M.,
  Kelley,J.C., Liu,L.-I., Marmaros,S.M., Merrick,J.M.,
  Moreno-Palancas,R.F., McDonald,L.A., Nguyen,D.T., Pelligrino,S.M.,
  Phillips,C.A., Ryder,S.E., Scott,J.L., Saudek,D.M., Shirley,R.,
  Small,K.V., Spriggs,I.A., Utterback,T.R., Weidman,J.F., Li,Y.,
  Bednarek,D.P., Cao,L., Cepeda,M.A., Coleman,T.A., Collins,E.J.,
  Dimke,D., Feng,D.-F., Ferrie,A., Fischer,C., Hastings,G.A., He,W.W.,
  Hu,J.S., Greene,J.M., Gruber,J., Hudson,P., Kim,A.K., Kozak,D.L.,
  Kunsch,C., Hungjun,J., Li,H., Weissner,P.S., Olsen,H., Raymond,L.,
  Wei,Y.F., Wing,J., Xu,C., Yu,G.L., Ruben,S.M., Dillion,P.J., Fannon
  ,M.R., Rosen,C.A., Haseltine,W.A., Fields,C., Fraser,C.M. and
  Venter,J.C.
  Initial assessment of human gene diversity and expression patterns
  based upon 83 million nucleotides of cDNA sequence
  Nature 377 (6547 Suppl), 3-174 (1995).
  95026280
  Bioinformatics
  The Institute for Genomic Research
  9712 Medical Center Drive, Rockville, MD 20850 USA
  Tel: 3018699056
  Fax: 3018699423
  Email: arkerlavet@tigr.org
  For clone availability, additional sequence and expression
  information related to this EST, please check the TIGR Human Gene
  Index (http://www.tigr.org/tadb/hgi/hgi.html)
  Seq primer: M13 Reverse.
  Location/Qualifiers
    1..171
    /organism="Homo sapiens"
    /db_xref="ATCC (inhost):179987"
    /db_xref="taxon:9606"
    /clone_lib="HSC172 cells II"
    /cell_type="fibroblast"
    /dev_stage="fetal"
    /note="Organ: lung; Vector: pBluescript SK-; Site_1: EcoRI
    ; Site_2: XhoI"
  61 a 34 c 48 g 28 t

BASE COUNT
ORIGIN
  Query Match 13.6%; Score 18; DB 9; Length 171;
  Best Local Similarity 100.0%; Pred. No. 84;
  Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy
  40 GATAAGGCCAAGCTGAAG 57
  |||||||
  61 GATAAGGCCAAGCTGAAG 78

Db
|||||
61 GATAAGGCCAAGCTGAAG 78

RESULT 35
AA692385
LOCUS
DEFINITION
  vt21a10.r1 Barstead mouse myotubes MPRB5 Mus musculus cDNA clone
  IMAGE:1163706 5' similar to gb:S54005 THYMOSIN BETA-10 (HUMAN);
  mRNA sequence.
ACCESSION
  AA692385
VERSION
  AA692385.1 GI:2693323
KEYWORDS
  EST.
  
```

```

W18665
LOCUS
DEFINITION
  mb98d04.r1 Soares mouse p3NMF19.5 Mus musculus cDNA clone
  IMAGE:337447 5' similar to gb:S54005 THYMOSIN BETA-10 (HUMAN);
  mRNA sequence.
ACCESSION
  W18665
VERSION
  W18665.1 GI:1294373
KEYWORDS
  EST.
SOURCE
  house mouse.
ORGANISM
  Mus musculus
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
  1 (bases 1 to 172)
  Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
  Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
  Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
  Theising,B., Wylie,T., Lennon,G., Soares,B., Willson,R. and
  Waterston,R.
  The WashU-HMI Mouse EST Project
  Unpublished (1996)
  Contact: Marra M/Mouse EST Project
  WashU-HMI Mouse EST Project
  Washington University School of Medicine
  4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
  Tel: 314 286 1800
  Fax: 314 286 1810
  Email: mouseest@watson.wustl.edu
  This clone is available royalty-free through LINL; contact the
  IMAGE Consortium (info@image.llnl.gov) for further information.
  MGI:218847
  Putative full length read
  Seq primer: mob.R5GA+ET
  High quality sequence stop: 160.
  Location/Qualifiers
    1..172
    /organism="Mus musculus"
    /db_xref="taxon:10090"
    /clone="IMAGE:337447"
    /clone_lib="Soares mouse, p3NMF19.5"
    /dev_stage="19.5 dpc total fetus"
    /lab_host="DH10B (ampicillin resistant)"
    /note="vector: pT7T3D (Pharmacia) with a modified
    polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
    was primed with a Not I - oligo(dT) primer [5'
    TGTTACCAATCTGAAGTGGAGCGCGCATTTTTTTTTTTTTTTT 3'],
    double-stranded cDNA was size selected, ligated to Eco RI
    adapters (Pharmacia), digested with Not I and cloned into
    the Not I and Eco RI sites of a modified pT7T3 vector
    (Pharmacia). Library went through one round of
    normalization to a Cot = 5. Library constructed by Bento
    Soares and M.Fatima Bonaldo. RNA was kindly provided by
    Dr. Minoru Ko (Wayne State University)."
  65 a 43 c 44 g 20 t

BASE COUNT
ORIGIN
  Query Match 13.6%; Score 18; DB 14; Length 172;
  Best Local Similarity 100.0%; Pred. No. 84;
  Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy
  40 GATAAGGCCAAGCTGAAG 57
  |||||||
  49 GATAAGGCCAAGCTGAAG 66

Db
|||||
49 GATAAGGCCAAGCTGAAG 66

RESULT 36
AA692385
LOCUS
DEFINITION
  vt21a10.r1 Barstead mouse myotubes MPRB5 Mus musculus cDNA clone
  IMAGE:1163706 5' similar to gb:S54005 THYMOSIN BETA-10 (HUMAN);
  mRNA sequence.
ACCESSION
  AA692385
VERSION
  AA692385.1 GI:2693323
KEYWORDS
  EST.
  
```

SOURCE house mouse.
ORGANISM Mus musculus
REFERENCE Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 173)
AUTHORS Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T., Geisler, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M., Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B., Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and Waterson, R.
TITLE The WashU-HMI Mouse EST Project
JOURNAL Unpublished (1996)
COMMENT Contact: Marra M/Mouse EST Project
WashU-HMI Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:629618
Seq primer: -28ml3 rev2 ET from Amersham
High quality sequence stop: 119.
FEATURES source
1..173
Location/Qualifiers
/organism="Mus musculus"
/strain="C3H"
/db_xref="taxon:10090"
/clone="IMAGE:1163706"
/clone_lib="Barstead mouse myotubes MPLRB5"
/cell_line="C2C12"
/lab_host="DH10B"
/note="Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker. Site.1: EcoRI; Site.2: NotI; 1st strand cDNA 3'; double-stranded cDNA was ligated to Eco RI adaptors [AATCGGATCTTG], digested with Not I and cloned into the Not I and Eco RI sites of the modified pT7T3 vector. Library constructed by Bob Barstead. The C2C12 cell line (available from ATCC, catalog # CRL-1772) differentiates rapidly, forming contractile myotubes and producing characteristic muscle proteins."
BASE COUNT 66 a 44 c 44 g 19 t
ORIGIN
Query Match 13.6%; Score 18; DB 9; Length 173;
Best Local Similarity 100.0%; Pred. No. 84; Mismatches 0; Indels 0; Gaps 0;
Matches 18; Conservative 0;
QY 40 GATAAGGCCAAGCTGAAG 57
|||||
Db 49 GATAAGGCCAAGCTGAAG 66
RESULT 37
T27818
LOCUS 174 bp mRNA linear EST 06-SEP-1995
DEFINITION EST17484 Human Endothelial cells Homo sapiens cDNA 5' end similar to thymosin beta-10 (H1:526), mRNA sequence.
T27818
ACCESSION T27818.1 GI:609916
VERSION
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 174)
REFERENCE Adams, M.D., Kervage, A.R., Fleischmann, R.D., Fuldner, R.A., Bult, C.J., Lee, N., Kirkness, E.F., Weinstock, K.G., Gocayne, J.D., White, O., Sutton, G., Blake, J.A., Brandon, R.C., Chiu, M.-W., Clayton, R.A., Cline, R.T., Cotton, M.D., Earle-Hughes, J., Fine, L.D., FitzGerald

L.M., FitzHugh, W.M., Fritchman, J.L., Geoghagen, N.S.M., Glodek, A., Gnehm, C.L., Hanna, M.C., Hedblom, E., Hinkle Jr, P.S., Kelley, J.M., Klimek, K.M., Kelley, J.C., Liu, L.-I., Marmaros, S.M., Merrick, J.M., Moreno-Palancas, R.F., McDonald, L.A., Nguyen, D.T., Pellegrino, S.M., Phillips, C.A., Ryder, S.E., Scott, J.L., Saudek, D.M., Shirley, R., Small, K.V., Spriggs, T.A., Utterback, T.R., Weidman, J.F., Li, F., Bednarek, D.P., Cao, L., Cepeda, M.A., Coleman, T.A., Collins, E.-J., Dinke, D., Feng, P., Ferrie, A., Fischer, C., Hastings, G.A., He, W.-W., Hu, J.-S., Greene, J.M., Gruber, J., Hudson, P., Kim, A., Kozak, D.L., Kunsch, C., Ji, H., Li, H., Meissner, P.S., Olsen, H., Raymond, L., Wei, Y.-F., Wang, J., Xu, C., Yu, G.-L., Ruben, S.M., Dillon, P.J., Fannon, M.R., Rosen, C.A., Haseltine, W.A., Fields, C., Fraser, C.M. and Venter, J.C.
Initial Assessment of Human Gene Diversity and Expression Patterns
Based Upon 83 Million Basepairs of cDNA Sequence
Nature 377, 3-174 (1995)
96026280
Other_ESTs: THC24270
Contact: Venter, JC
The Institute for Genomic Research
932 Clopper Rd, Gaithersburg, MD 20878
Tel: 3018699056
Fax: 3018699423
Email: tdbinfo@tdb.tigr.org
For clone availability, additional sequence and expression information related to this EST, please contact the TIGR Database (tdbinfo@tdb.tigr.org)
Seq primer: M13 Reverse.
FEATURES source
1..174
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="ATCC (inhost):100256"
/db_xref="taxon:9606"
/clone_lib="Human Endothelial cells"
/tissue_type="endothelial"
/note="Organ: colon"
BASE COUNT 62 a 37 c 52 g 23 t
ORIGIN
Query Match 13.6%; Score 18; DB 14; Length 174;
Best Local Similarity 100.0%; Pred. No. 85; Mismatches 0; Indels 0; Gaps 0;
Matches 18; Conservative 0;
QY 40 GATAAGGCCAAGCTGAAG 57
|||||
Db 78 GATAAGGCCAAGCTGAAG 95
RESULT 38
BE774172/c
LOCUS 179 bp mRNA linear EST 20-SEP-2000
DEFINITION MR1-UM0009-300500-017-e07 UM0009 Homo sapiens cDNA, mRNA sequence.
ACCESSION BE774172
VERSION BE774172.1 GI:10227879
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 179)
REFERENCE Dias Neto, E., Garcia Correa, R., Verjowski-Almeida, S., Briones, M.R., Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F., Goldman, G.H., Carvalho, A.F., Matsukuma, A., Bala, G.S., Simpson, D.H., Brunstein, A., deOliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and Simpson, A.J.
Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
20202663
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research

Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
 Tel: +55-11-2704922
 Fax: +55-11-2707001
 Email: asimpson@ludwig.org.br
 This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
 (http://www.ludwig.org.br/scripts/gethtml2.pl?tl=st2-MR1-UM0009-300
 500-017-e07&t3=2000-05-30&t4=1)
 Seq primer: puc 18 forward
 High quality sequence stop: 179.

FEATURES

Source

1. 179
 Location/Qualifiers
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_lib="UM0009"
 /dev_stage="Adult"
 /note="Organ: uterus; Vector: puc18; Site_1: SmaI; Site_2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196 716 - Ludwig Institute for Cancer Research) profiles into the puc 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

29 a 47 c 46 g 57 t

BASE COUNT

ORIGIN

Query Match 13.6%; Score 18; DB 12; Length 179;
 Best Local Similarity 100.0%; Pred. No. 85;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 40 GATAAGGCCAAGCTGAAG 57

Db 148 GATAAGGCCAAGCTGAAG 131

RESULT 39

AA068081

LOCUS

DEFINITION

mm56e02.r1 Stratagene mouse embryonic carcinoma (#937317) Mus

musculus cDNA clone IMAGE:532442 5' similar to gb:S54005 THYMOSIN

BETA-10 (HUMAN); mRNA sequence.

AA068081.1 GI:1566362

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

1. 180
 Location/Qualifiers
 /organism="Mus musculus"
 /cultivar="3T3"
 /db_xref="taxon:10090"
 /clone="R3DAB96"
 /clone_lib="R3DA"
 /tissue_type="adipose"
 /note="Site_1: EcoRI; Site_2: XhoI"
 Unpublished (1994)
 Contact: Marra M/Mouse EST Project
 WasRU-HMI Mouse EST Project
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810

Email: mouseest@watson.wustl.edu
 This clone is available royalty-free through LLNL; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 MGI:319378

Seq primer: -28ml3 rev1 ET from Amersham

High quality sequence stop: 50.

Location/Qualifiers

1. 180

/organism="Mus musculus"

/db_xref="taxon:10090"

FEATURES

Source

/clone="IMAGE:532442"
 /clone_lib="Stratagene mouse embryonic carcinoma (#937317)"
 /tissue_type="carcinoma"
 /dev_stage="embryonic"
 /lab_host="SOLR (kanamycin resistant)"
 /note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2: XhoI; Cloned unidirectionally. Primer: Oligo dt. P19 cell line. Average insert size: 1.0 kb; Uni-ZAP XR vector; -5' adaptor sequence: 5' GAATTCGACGAG 3' -3' adaptor sequence: 5' CTCGAGTTTTTTTTTTTTTTT 3'"

68 a 38 c 52 g 22 t

BASE COUNT

ORIGIN

Query Match 13.6%; Score 18; DB 9; Length 180;
 Best Local Similarity 100.0%; Pred. No. 86;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 40 GATAAGGCCAAGCTGAAG 57

Db 69 GATAAGGCCAAGCTGAAG 86

RESULT 40

AW647096

LOCUS

DEFINITION

EST325639 R3DA Mus musculus cDNA clone R3DAB96, mRNA sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

1. 180
 Location/Qualifiers
 /organism="Mus musculus"
 /cultivar="3T3"
 /db_xref="taxon:10090"
 /clone="R3DAB96"
 /clone_lib="R3DA"
 /tissue_type="adipose"
 /note="Site_1: EcoRI; Site_2: XhoI"
 Unpublished (1994)
 Other ESTs: TC84502
 Contact: Julie Earle-Hughes
 TIGR
 The Institute For Genomic Research
 9712 Medical Center Drive, Rockville, MD 20850, USA
 Tel: 301 838 0200
 Fax: 301 838 0208.

FEATURES

Source

1. 180
 Location/Qualifiers
 /organism="Mus musculus"
 /cultivar="3T3"
 /db_xref="taxon:10090"
 /clone="R3DAB96"
 /clone_lib="R3DA"
 /tissue_type="adipose"
 /note="Site_1: EcoRI; Site_2: XhoI"
 Unpublished (1994)
 Other ESTs: TC84502
 Contact: Julie Earle-Hughes
 TIGR
 The Institute For Genomic Research
 9712 Medical Center Drive, Rockville, MD 20850, USA
 Tel: 301 838 0200
 Fax: 301 838 0208.

BASE COUNT

ORIGIN

Query Match 13.6%; Score 18; DB 10; Length 180;
 Best Local Similarity 100.0%; Pred. No. 86;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 40 GATAAGGCCAAGCTGAAG 57

Db 69 GATAAGGCCAAGCTGAAG 86

RESULT 41

AA666472

LOCUS

DEFINITION

AA666472 181 bp mRNA linear EST 19-NOV-1997

vm48b04.r1 Stratagene mouse Tcell 937311 Mus musculus cDNA clone

IMAGE:1001455 5' similar to gb:S54005 THYMOSIN BETA-10 (HUMAN);

mRNA sequence.

ACCESSION
AA666472
VERSION
AA666472.1 GI:2625173
KEYWORDS
EST.
SOURCE
house mouse.
ORGANISM
Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 181)

Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T.,
Giesel, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M.,
Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B.,
Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and
Waterson, R.

The WashU-HMI Mouse EST Project

Unpublished (1996)

Contact: Marra M/Mouse EST Project

WashU-HMI Mouse EST Project

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

Tel: 314 286 1800

Fax: 314 286 1810

Email: mouseest@watson.wustl.edu

This clone is available royalty-free through LLNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.

MGI:565671

Seq primer: -28mi3 rev1 ET from Amersham.

Location/Qualifiers

1. .181

/organism="Mus musculus"

/db_xref="taxon:10090"

/clone="IMAGE:1001455"

/clone_lib="Stratagene mouse Tcell 937311"

/tissue.type="Tcell"

/dev.stage="M30 CD4+ cells"

/lab_host="SOLR (kanamycin resistant)"

/note="Organ: blood; Vector: pBluescript SK-; Site_1:

EcoRI; Site_2: XhoI; Cloned unidirectionally. Primer:

Oligo dt. M30 CD4+ cells. Average insert size: 1.0 kb;

Uni-ZAP XR Vector; -5' adaptor sequence: 5' GAATTCGGCAGCAG

3' -3' adaptor sequence: 5' CTCGAGTTTTTTTTTTTTTTT 3'

20 a 52 c 40 g 69 t

BASE COUNT
ORIGIN

Query Match 13.6%; Score 18; DB 9; Length 181;

Best Local Similarity 100.0%; Pred. No. 86;

Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 40 GATAAGGCCAAGCTGAAG 57

|||||

Db 103 GATAAGGCCAAGCTGAAG 86

RESULT 42

AA316715

LOCUS

DEFINITION

EST188545 HCC cell line (metastasis to liver in mouse) II Homo

sapiens cDNA 5' end similar to similar to thymosin beta-10

(GB:S54005), mRNA sequence.

ACCESSION

AA316715

VERSION

AA316715.1 GI:1969214

KEYWORDS

EST.

SOURCE

human.

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

AUTHORS

Adams, M.D., Kerlavage, A.R., Fleischmann, R.D., Fuldner, R.A., Bult

, C.J., Lee, N.H., Kirkness, E.F., Weinstock, K.G., Gocayne, J.D., White

, O., Sutton, G., Blake, J.A., Brandon, R.C., Man-Wai, C., Clayton, R.A.,

, L.M., Fitchugh, W.M., Fritchman, J.L., Geoghagen, N.S., Glodek, A.,

Gnehm, C.L., Hanna, M.C., Hedblom, E., Hinkley, P.S.Jr., Kelley, J.M.,

Kelley, J.C., Liu, L.-I., Marmaros, S.M., Merrick, J.M.,
Moreno-Palannes, R.F., McDonald, J.A., Nguyen, D.T., Pelligrino, S.M.,
Phillips, C.A., Ryder, S.E., Scott, J.L., Saudek, D.M., Shirley, R.,
Small, K.V., Spriggs, T.A., Utterback, T.R., Weidman, J.F., Li, Y.,
Bednarik, P., Cao, L., Cepeda, M.A., Coleman, T.A., Collins, E.J.,
Dimke, D., Feng, D.-F., Ferrie, A., Fischer, C., Hastings, G.A., He, W.W.,
Hu, J.S., Greene, J.M., Gruber, J., Hudson, P., Kim, A.K., Kozak, D.L.,
Kunsch, C., Hungjun, J., Li, H., Meissner, P.S., Olsen, H., Raymond, L.,
Wei, Y.F., Wing, J., Xu, C., Yu, G.L., Ruben, S.M., Dillion, P.J., Fannon
M.R., Rosen, C.A., Haseltine, W.A., Fields, C., Fraser, C.M. and
Venter, J.C.

Initial assessment of human gene diversity and expression patterns
based upon 83 million nucleotides of cDNA sequence
Nature 377 (6547 Suppl.), 3-174 (1995)

TITLE
JOURNAL
MEDLINE
COMMENT

96026280

Bioinformatics

Contact: Kerlavage, AR

The Institute for Genomic Research

9712 Medical Center Drive, Rockville, MD 20850 USA

Tel: 3018699056

Fax: 3018699423

Email: arkerlav@tigr.org

For clone availability, additional sequence and expression

information related to this EST, please check the TIGR Human Gene

Index (<http://www.tigr.org/tdb/tgi/hgi.html>)

Seq primer: M3 Reverse.

Location/Qualifiers

1. .182

/organism="Homo sapiens"

/db_xref="ATCC (inhost):113617"

/db_xref="taxon:9606"

/clone_lib="HCC cell line (metastasis to liver in mouse)

II"

/tissue.type="colon"

/cell_type="KM12SM"

/cell_line="KM12C(HCC)metastasis into mouse (liver)"

/note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2:

XhoI"

63 a 39 c 53 g 26 t 1 others

BASE COUNT

ORIGIN

Query Match 13.6%; Score 18; DB 9; Length 182;

Best Local Similarity 100.0%; Pred. No. 86;

Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 40 GATAAGGCCAAGCTGAAG 57

|||||

Db 80 GATAAGGCCAAGCTGAAG 97

RESULT 43

N83763

LOCUS

DEFINITION

clone KK2711F Human fetal heart, Lambda ZAP Express Homo sapiens cDNA

clone KK2711 5' similar to THYMOSIN BETA-10, mRNA sequence.

ACCESSION

N83763

VERSION

N83763.1 GI:1259388

KEYWORDS

EST.

SOURCE

human.

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

AUTHORS

Liew, C.C.

CDNAs from fetal heart (1996)

1 (bases 1 to 184)

TITLE

Unpublished (1996)

CONTACT: Liew CC

Brigham and Women's Hospital

Harvard Medical School

75 Francis St. Boston, MA 02115, USA

Tel: 6177328915

Fax: 6179750995

Email: clliew@rics.bwh.harvard.edu

Seq primer: GAATTAACCTCCTACTAAAGG.

FEATURES

source

Location/Qualifiers
1. .184
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="KK2711"

/clone_lib="Human fetal heart, Lambda ZAP Express"
/lab_host="E. coli XL1-Blue"
/note="Vector: Lambda ZAP Express; Site_1: EcoRI; Site_2: XhoI; mRNA was purified from human fetal hearts (8-10 weeks). cDNA was synthesized using a XhoI-Oligo dT adaptor-primer. EcoRI adaptors were ligated, followed by digestion with XhoI, for directional cloning into predigested lambda ZAP Express."

63 a 38 c 57 g 26 t

BASE COUNT

ORIGIN

Query Match 13.6%; Score 18; DB 14; Length 184;
Best Local Similarity 100.0%; Pred. No. 86;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 40 GATAAGGCCCAAGCTGAAG 57

Db 94 GATAAGGCCCAAGCTGAAG 111

RESULT 44

AW211038

LOCUS

DEFINITION AW211038 187 bp mRNA linear EST 03-DEC-1999
similar to gb:S54005 THYMOSIN BETA-10 (HUMAN);, mRNA sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.

REFERENCE

AUTHORS

TITLE

COMMENT

National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Lohar Hennighausen Ph.D., Chu-Xia Deng Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html

MGI:1030265

Seq primer: -40RP from Gibco

High quality sequence stop: 181.

FEATURES

source

Location/Qualifiers
1. .187
/organism="Mus musculus"
/strain="129,C57BL/6J,FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:2649813"
/clone_lib="NCI-CGAP_Mam3"
/tissue_type="tumor, gross tissue"
/dev_stage="10 months"
/lab_host="DH10B"

/note="Organ: mammary; Vector: pCMV-SPORT6; Site_1: SalI;
Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT.
Library constructed by Life Technologies. Investigators
providing samples: Lohar Hennighausen/Chu-Xia Deng, NIH
Reference for transgenic model: Xu et al., Nature Genetics
22, 37-43 (1999)."

68 a 41 c 57 g 21 t

BASE COUNT

ORIGIN

Query Match 13.6%; Score 18; DB 10; Length 187;
Best Local Similarity 100.0%; Pred. No. 87;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 40 GATAAGGCCCAAGCTGAAG 57

Db 93 GATAAGGCCCAAGCTGAAG 110

RESULT 45

AA072898

LOCUS

DEFINITION AA072898 189 bp mRNA linear EST 07-FEB-1997
mm72d12.t1 Stratagene mouse macrophage (#937306) Mus musculus cDNA
clone IMAGE:533975 5' similar to gb:S54005 THYMOSIN BETA-10 (HUMAN
);, mRNA sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Mus musculus

house mouse

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 189)

Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,

Geisel,S., Kucaba,T., Lacy,M., Le.M., Martin,J., Morris,M.,

Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,

Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and

Waterston,R.

The WashU-HMNI Mouse EST Project

Unpublished (1996)

Contact: Marra M/Mouse EST Project

WashU-HMNI Mouse EST Project

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

Tel: 314 286 1800

Fax: 314 286 1810

Email: mouseest@watson.wustl.edu

This clone is available royalty-free through LLNL; contact the

IMAGE Consortium (info@image.llnl.gov) for further information.

MGI:320911

Trace considered overall poor quality

Seq primer: -28ml3 revl ET from Amersham

High quality sequence stop: 1.

Location/Qualifiers

1. .189

/organism="Mus musculus"

/db_xref="taxon:10090"

/clone="IMAGE:533975"

/clone_lib="Stratagene mouse macrophage (#937306)"

/tissue_type="macrophage"

/dev_stage="WEHI-3 cell line"

/lab_host="SOLR (kanamycin resistant)"

/note="Organ: blood; Vector: pBluescript SK-; Site_1:

EcoRI; Site_2: XhoI; Cloned unidirectionally. Primer:

Oligo dT. WEHI-3 cell line. Average insert size: 1.5 kb;

Uni-ZAP XR Vector; ~5' adaptor sequence: 5' GAATTCGGCAGG

3' -3' adaptor sequence: 5' CTCGAGTTTTTTTTTTTTTTT 3' "

BASE COUNT 69 a 44 c 53 g 23 t

ORIGIN

Query Match 13.6%; Score 18; DB 9; Length 189;

Best Local Similarity 100.0%; Pred. No. 87;

Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 40 GATAAGGCCCAAGCTGAAG 57

Db 69 GATAAGGCCCAAGCTGAAG 86

RESULT 46

AA087727

LOCUS
DEFINITION AA087727 195 bp mRNA linear EST 23-OCT-1996
mol3e09.r1 Life Tech mouse embryo 10 5dpc 10665016 Mus musculus
cDNA clone IMAGE:553480 5' similar to gb:S54005 THYMOSIN BETA-10
(HUMAN);, mRNA sequence.

ACCESSION
VERSION AA087727
KEYWORDS AA087727.1 GI:1630936
SOURCE EST.
ORGANISM house mouse.

REFERENCE
AUTHORS 1 (bases 1 to 195)
Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
Waterston,R.

TITLE The WashU-HMI Mouse EST Project
JOURNAL Unpublished (1996)
COMMENT Contact: Marra M/Mouse EST Project
WashU-HMI Mouse EST Project
Washington University School of MedicineP
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@wustl.edu
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:334272
Seq primer: -28M13 rev1 from Amersham
High quality sequence stop: 163.
Location/Qualifiers
1..195
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="IMAGE:553480"
/clone_lib="Life Tech mouse embryo 10 5dpc 10665016"
/tissue_type="embryo"
/dev_stage="10.5dpc embryos"
/lab_host="DH10B"
/note="Organ: whole embryo; Vector: PCMV-SPORT2; Site_1:
Salt; Site_2: NotI; Cloned unidirectionally. Primer:
Oligo dt. 10.5dpc embryos. PCMV-SPORT2 vector."
71 a 43 c 57 g 24 t

BASE COUNT
ORIGIN
Query Match 13.6%; Score 18; DB 9; Length 195;
Best Local Similarity 100.0%; Pred. No. 88;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 40 GATAAGGCCAAGCTGAAG 57
|||||
Db 84 GATAAGGCCAAGCTGAAG 101
|||||

RESULT 47
AA454807
LOCUS AA454807 195 bp mRNA linear EST 06-JUN-1997
DEFINITION zx77f10.r1 Soares ovary tumor NbHOT Homo sapiens cDNA clone
IMAGE:809803 5' similar to gb:S54005 THYMOSIN BETA-10 (HUMAN);,
mRNA sequence.

ACCESSION
VERSION AA454807.1 GI:2177583
KEYWORDS EST.
SOURCE human.

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 195)
Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S.,
Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M., Martin,J., Moore,B.,
Schellenberg,K., Steptoe,M., Tan,F., Theising,B., White,Y., Wylie

TITLE
JOURNAL
COMMENT 'T., Waterston,R. and Wilson,R.
WashU-Merck EST Project 1997
Unpublished (1997)
Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@wustl.edu
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Trace considered overall poor quality
Seq primer: -28M13 rev2 ET from Amersham
High quality sequence stop: 1.
Location/Qualifiers
1..195
/organism="Homo sapiens"
/db_xref="GDB:6039563"
/db_xref="taxon:9606"
/clone="IMAGE:809803"
/clone_lib="Soares ovary tumor NbHOT"
/sex="Female"
/tissue_type="ovarian tumor"
/lab_host="DH10B (ampicillin resistant)"
/note="Organ: ovary; Vector: pT7T3D (Pharmacia) with a
modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
strand cDNA was primed with a Not I - oligo(dt) primer [5'
TGTTACCAATCTGAAGTGGAGCGCGCGTGTGTGTGTGTGTGTGT 3'],
double-stranded cDNA was size selected, ligated to Eco RI
adapters (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of a modified pT7T3 vector
(Pharmacia). Library constructed by Bento Soares and
M.Patima Bonaldo."
64 a 45 c 53 g 33 t

BASE COUNT
ORIGIN
Query Match 13.6%; Score 18; DB 9; Length 195;
Best Local Similarity 100.0%; Pred. No. 88;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 40 GATAAGGCCAAGCTGAAG 57
|||||
Db 73 GATAAGGCCAAGCTGAAG 90
|||||

RESULT 48
AV044717/c
LOCUS AV044717 198 bp mRNA linear EST 23-NOV-1999
DEFINITION AV044717 Mus musculus adult C57BL/6J testis Mus musculus cDNA clone
1700038G19, mRNA sequence.

ACCESSION
VERSION AV044717
KEYWORDS AV044717.2 GI:4864382
SOURCE EST.
house mouse.
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 198)
Carninci,P., Shibata,K., Ozawa,Y., Konno,H., Itoh,M., Aizawa,K.,
Akahira,S., Akiyama,J., Fukuda,S., Fukunishi,Y., Funayama,T., Hara
A., Hayatsu,N., Hori,F., Ishikawa,T., Itoh,M., Izawa,M., Kawai,J.,
Kikuchi,N., Kojima,Y., Matsuyama,T., Nitsuma,H., Oda,H., Owa,C.,
Sato,K., Shibata,Y., Shigemoto,Y., Shiraki,T., Sogabe,Y., Sugahara
Y., Suzuki,H., Suzuki,H., Tateno,M., Tomaru,Y., Tomioka,N.,
Watanabe,S., Yagame,M., Yamamura,T., Yokota,T., Yoshino,M.,
Muramatsu,M., Okazaki,Y. and Hayashizaki,Y.

RIKEN Mouse ESTs
Unpublished (1999)
Contact: Chie Owa
Genome Science Laboratory
RIKEN
3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan
Tel: 81-298-36-9145

Information related to this EST, please check the TIGR Human Gene
Index (<http://www.tigr.org/tdb/hgi/hgi.html>)

Seq primer: M13 Reverse.

FEATURES
source

Location/Qualifiers

1..201
/organism="Homo sapiens"
/db_xref="ATCC (inhost):114687"
/db_xref="taxon:9606"
/clone_lib="Aorta endothelial cells"
/cell_type="endothelial cell"
/dev_stage="adult"

/note="Organ: aorta; Vector: pBluescript SK-; Site_1:
EcoRI; Site_2: XhoI"

BASE COUNT 66 a 45 c 57 g 30 t 3 others

ORIGIN

Query Match 13.6%; Score 18; DB 9; Length 201;

Best Local Similarity 100.0%; Pred. No. 89;

Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 40 GATAAGGCCAAGCTGAAG 57

|||||

Db 100 GATAAGGCCAAGCTGAAG 117

Search completed: June 2, 2003, 16:39:10

Job time : 1516 secs

THIS PAGE BLANK (CONT)

GenCore version 5.1.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 2, 2003, 15:41:34 ; Search time 66 seconds
(without alignments)
613.354 Million cell updates/sec

Title: US-09-915-178-1

Perfect score: 132

Sequence: 1 atggcacacaaactagacct.....atggagtgaaatttctga 132

Scoring table: OLIGO_NUC

Gapop 60.0 , Gapext 60.0

Searched: 441362 seqs, 153338381 residues

Word size : 10

Total number of hits satisfying chosen parameters: 34415

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database :

Issued_Patents_NA.*
1: /cgn2.6/ptodata/1/ina/5A_COMB.seq.*
2: /cgn2.6/ptodata/1/ina/5B_COMB.seq.*
3: /cgn2.6/ptodata/1/ina/6A_COMB.seq.*
4: /cgn2.6/ptodata/1/ina/6B_COMB.seq.*
5: /cgn2.6/ptodata/1/ina/PTUS_COMB.seq.*
6: /cgn2.6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
c 1	17	12.9	439	4	US-09-397-787-308
c 2	17	12.9	776	4	US-09-484-970B-82
c 3	17	12.9	2090	2	US-08-445-520B-8
4	17	12.9	2090	3	US-08-451-946B-7
5	17	12.9	2090	3	US-08-446-938B-7
6	17	12.9	2090	3	US-08-311-703A-7
7	17	12.9	2090	3	US-08-446-939B-7
8	17	12.9	2090	3	US-09-183-543-7
9	17	12.9	2090	4	US-08-446-936A-7
10	17	12.9	2090	4	US-09-239-864A-10
11	17	12.9	2090	5	PCT-US92-09326-3
12	17	12.9	2095	1	US-08-361-873A-1
13	17	12.9	5057	2	US-08-365-486A-12
c 14	17	12.9	5057	4	US-08-880-342-12
c 15	17	12.9	5108	1	US-07-642-002-1
c 16	16	12.1	220	2	US-08-700-637-9
17	16	12.1	1232	4	US-09-392-184-4
18	16	12.1	3437	3	US-08-704-711A-9
19	16	12.1	3437	4	US-09-521-220-9
20	16	12.1	3456	3	US-08-704-711A-8
21	16	12.1	3456	4	US-09-521-220-8
c 22	16	12.1	4380	1	US-07-582-945-1
c 23	16	12.1	4380	2	US-08-453-141-1
c 24	16	12.1	4380	3	US-08-293-314-1
c 25	16	12.1	8201	1	US-08-253-155A-9
c 26	16	12.1	12848	4	US-09-453-702B-252
27	16	12.1	17056	4	US-09-245-041-3

16	12.1	38155	4	US-09-453-702B-79	Sequence 79, Appl	
16	12.1	48908	4	US-09-453-702B-137	Sequence 137, App	
30	11.4	75	1	US-08-153-799-9	Sequence 9, Appl	
31	11.4	460	4	US-09-040-984-83	Sequence 83, Appl	
32	11.4	460	4	US-09-123-912-83	Sequence 83, Appl	
33	11.4	460	4	US-09-643-597-83	Sequence 83, Appl	
c 34	11.4	975	1	US-08-671-525B-9	Sequence 9, Appl	
c 35	11.4	975	1	US-08-672-109B-9	Sequence 9, Appl	
c 36	11.4	975	1	US-08-842-045-9	Sequence 9, Appl	
c 37	11.4	975	2	US-08-842-238-9	Sequence 9, Appl	
c 38	11.4	978	3	US-08-629-335B-9	Sequence 9, Appl	
c 39	11.4	978	3	US-08-706-281A-17	Sequence 17, Appl	
c 40	11.4	978	4	US-09-097-231-17	Sequence 17, Appl	
41	11.4	1586	4	US-09-152-060-12	Sequence 12, Appl	
42	11.4	1711	4	US-09-453-702B-105	Sequence 105, App	
43	11.4	1897	1	US-08-615-170-18	Sequence 18, Appl	
c 44	11.4	1964	4	US-09-434-613-2	Sequence 2, Appl	
45	11.4	2012	4	US-09-149-476-132	Sequence 132, App	
46	11.4	3102	4	US-09-336-643A-17	Sequence 17, Appl	
47	11.4	3876	1	US-08-494-714-1	Sequence 1, Appl	
48	11.4	3876	5	PCT-US96-10782-1	Sequence 1, Appl	
49	11.4	4350	4	US-09-295-593-37	Sequence 37, Appl	
c 50	11.4	4833	4	US-09-066-047-1	Sequence 1, Appl	
c 51	11.4	5541	1	US-08-920-812-20	Sequence 20, Appl	
c 52	11.4	5541	1	US-08-920-827-20	Sequence 20, Appl	
c 53	11.4	5541	1	US-08-921-177-20	Sequence 20, Appl	
c 54	11.4	5541	1	US-08-362-577C-20	Sequence 20, Appl	
c 55	11.4	5541	2	US-08-920-828-20	Sequence 20, Appl	
c 56	11.4	7676	1	US-08-451-777A-7	Sequence 7, Appl	
c 57	11.4	7676	2	US-08-451-778A-7	Sequence 7, Appl	
c 58	11.4	7676	2	US-08-998-208-7	Sequence 7, Appl	
c 59	11.4	7676	5	PCT-US95-06743-7	Sequence 7, Appl	
60	11.4	8367	2	US-08-583-562B-7	Sequence 7, Appl	
61	11.4	8367	2	US-08-779-113-7	Sequence 7, Appl	
62	11.4	10718	3	US-08-325-426B-1	Sequence 1, Appl	
63	11.4	12666	4	US-08-961-527-137	Sequence 137, App	
64	11.4	19736	4	US-09-740-035-3	Sequence 3, Appl	
c 65	11.4	4403765	4	US-09-103-840A-2	Sequence 2, Appl	
c 66	11.4	4411529	4	US-09-103-840A-1	Sequence 1, Appl	
c 67	14	10.6	30	US-08-612-973-34	Sequence 34, Appl	
c 68	14	10.6	30	US-08-927-597-34	Sequence 34, Appl	
c 69	14	10.6	50	US-08-171-389-426	Sequence 426, App	
c 70	14	10.6	50	US-08-123-936-426	Sequence 426, App	
c 71	14	10.6	50	US-08-475-228A-426	Sequence 426, App	
c 72	14	10.6	50	US-08-482-080A-426	Sequence 426, App	
c 73	14	10.6	50	US-09-354-947-426	Sequence 426, App	
c 74	14	10.6	50	PCT-US93-12388-426	Sequence 426, App	
c 75	14	10.6	51	US-08-171-389-427	Sequence 427, App	
c 76	14	10.6	51	US-08-123-936-427	Sequence 427, App	
c 77	14	10.6	51	US-08-475-228A-427	Sequence 427, App	
c 78	14	10.6	51	US-08-482-080A-427	Sequence 427, App	
c 79	14	10.6	51	US-09-354-947-427	Sequence 427, App	
c 80	14	10.6	51	PCT-US93-12388-427	Sequence 427, App	
c 81	14	10.6	213	3	US-08-513-974B-37	Sequence 37, Appl
c 82	14	10.6	233	4	US-09-085-199B-24	Sequence 24, Appl
c 83	14	10.6	240	3	US-08-513-974B-318	Sequence 318, App
c 84	14	10.6	339	1	US-08-171-718-118	Sequence 118, App
c 85	14	10.6	339	3	US-08-478-087-118	Sequence 118, App
c 86	14	10.6	372	3	US-09-188-930-60	Sequence 60, Appl
c 87	14	10.6	531	4	US-09-404-879A-138	Sequence 138, App
c 88	14	10.6	571	4	US-09-404-879A-106	Sequence 106, App
c 89	14	10.6	737	1	US-08-463-115-19	Sequence 19, Appl
90	14	10.6	737	1	US-08-465-388-19	Sequence 19, Appl
c 91	14	10.6	798	4	US-08-998-416-463	Sequence 463, App
c 92	14	10.6	858	4	US-09-334-938-12	Sequence 12, Appl
93	14	10.6	976	1	US-08-266-451B-25	Sequence 25, Appl
c 94	14	10.6	976	2	US-08-748-725-25	Sequence 25, Appl
c 95	14	10.6	1033	4	US-09-171-209-73	Sequence 73, Appl
c 96	14	10.6	1089	4	US-09-513-838-1	Sequence 1, Appl
c 97	14	10.6	1125	3	US-08-591-685-10	Sequence 10, Appl
c 98	14	10.6	1137	4	US-09-045-583-6	Sequence 6, Appl
c 99	14	10.6	1137	4	US-09-534-185-6	Sequence 6, Appl
c 100	14	10.6	1292	4	US-09-198-603C-7	Sequence 7, Appl

101	14	10.6	1336	1	US-08-107-684B-1	Sequence 1, Appl	174	14	10.6	2074	3	US-09-248-137-11	Sequence 11, Appl
102	14	10.6	1337	2	US-08-743-637B-32	Sequence 32, Appl	175	14	10.6	2190	3	US-08-508-761B-5	Sequence 5, Appl
103	14	10.6	1337	3	US-08-526-840B-32	Sequence 32, Appl	176	14	10.6	2241	2	US-08-838-219B-20	Sequence 20, Appl
104	14	10.6	1359	3	US-08-227-496C-17	Sequence 17, Appl	177	14	10.6	2241	3	US-09-233-336A-20	Sequence 20, Appl
105	14	10.6	1364	4	US-08-763-704A-2	Sequence 2, Appl	178	14	10.6	2241	3	US-09-233-752A-20	Sequence 20, Appl
106	14	10.6	1442	1	US-08-247-908A-1	Sequence 1, Appl	179	14	10.6	2241	4	US-09-402-036-20	Sequence 20, Appl
107	14	10.6	1442	1	US-08-453-942-1	Sequence 1, Appl	180	14	10.6	2241	4	US-09-904-226-20	Sequence 20, Appl
108	14	10.6	1442	1	US-08-926-885A-1	Sequence 1, Appl	181	14	10.6	2257	1	US-08-171-718-15	Sequence 15, Appl
109	14	10.6	1442	5	PCT-US94-05290-1	Sequence 1, Appl	182	14	10.6	2257	3	US-08-478-087-15	Sequence 15, Appl
110	14	10.6	1443	1	US-08-443-965B-10	Sequence 10, Appl	183	14	10.6	2328	4	US-09-513-838-5	Sequence 5, Appl
111	14	10.6	1443	2	US-08-423-989B-10	Sequence 10, Appl	184	14	10.6	2370	2	US-08-838-219B-19	Sequence 19, Appl
112	14	10.6	1443	2	US-08-443-966B-10	Sequence 10, Appl	185	14	10.6	2370	3	US-09-233-336A-19	Sequence 19, Appl
113	14	10.6	1452	4	US-09-039-982A-36	Sequence 36, Appl	186	14	10.6	2370	3	US-09-233-752A-19	Sequence 19, Appl
114	14	10.6	1452	4	US-09-039-641-36	Sequence 36, Appl	187	14	10.6	2370	4	US-09-402-036-19	Sequence 19, Appl
115	14	10.6	1452	4	US-09-039-762A-36	Sequence 36, Appl	188	14	10.6	2370	4	US-09-904-226-19	Sequence 19, Appl
116	14	10.6	1452	4	US-09-042-492D-36	Sequence 36, Appl	189	14	10.6	2403	1	US-08-471-033-30	Sequence 30, Appl
117	14	10.6	1452	4	US-08-913-612A-36	Sequence 36, Appl	190	14	10.6	2403	1	US-08-471-044-30	Sequence 30, Appl
118	14	10.6	1515	4	US-09-240-915-4	Sequence 4, Appl	191	14	10.6	2403	2	US-08-463-483A-30	Sequence 30, Appl
119	14	10.6	1515	4	US-09-240-915-5	Sequence 5, Appl	192	14	10.6	2403	2	US-08-471-046A-30	Sequence 30, Appl
120	14	10.6	1515	4	US-09-591-435-4	Sequence 4, Appl	193	14	10.6	2403	2	US-08-470-566B-30	Sequence 30, Appl
121	14	10.6	1515	4	US-09-591-435-5	Sequence 5, Appl	194	14	10.6	2403	2	US-08-838-219B-7	Sequence 7, Appl
122	14	10.6	1518	4	US-09-240-915-1	Sequence 1, Appl	195	14	10.6	2403	2	US-08-469-334-30	Sequence 30, Appl
123	14	10.6	1518	4	US-09-240-915-2	Sequence 2, Appl	196	14	10.6	2403	3	US-09-300-529-30	Sequence 30, Appl
124	14	10.6	1518	4	US-09-591-435-1	Sequence 1, Appl	197	14	10.6	2403	3	US-09-233-336A-7	Sequence 7, Appl
125	14	10.6	1518	4	US-09-591-435-2	Sequence 2, Appl	198	14	10.6	2403	3	US-09-233-752A-7	Sequence 7, Appl
126	14	10.6	1584	1	US-08-247-908A-10	Sequence 10, Appl	199	14	10.6	2403	4	US-09-402-036-7	Sequence 7, Appl
127	14	10.6	1584	1	US-08-453-942-10	Sequence 10, Appl	200	14	10.6	2502	1	US-09-904-228-7	Sequence 7, Appl
128	14	10.6	1584	2	US-08-926-885A-10	Sequence 10, Appl	201	14	10.6	2502	1	US-08-062-472B-2	Sequence 2, Appl
129	14	10.6	1584	5	PCT-US94-05290-10	Sequence 10, Appl	202	14	10.6	2514	4	US-09-144-914-3	Sequence 3, Appl
130	14	10.6	1599	1	US-08-196-003-2	Sequence 2, Appl	203	14	10.6	2574	4	US-09-255-829-21	Sequence 21, Appl
131	14	10.6	1599	2	US-08-688-870-1	Sequence 1, Appl	204	14	10.6	2574	4	US-09-255-829-27	Sequence 27, Appl
132	14	10.6	1599	2	US-08-688-870-1	Sequence 1, Appl	205	14	10.6	2574	4	US-09-255-829-28	Sequence 28, Appl
133	14	10.6	1599	3	US-08-933-824-2	Sequence 2, Appl	206	14	10.6	2700	1	US-08-484-105-5	Sequence 5, Appl
134	14	10.6	1599	4	US-09-264-465-2	Sequence 2, Appl	207	14	10.6	2700	1	US-08-484-106-5	Sequence 5, Appl
135	14	10.6	1611	4	US-09-039-984A-35	Sequence 35, Appl	208	14	10.6	2992	3	US-09-362-123A-3	Sequence 3, Appl
136	14	10.6	1611	4	US-09-039-984A-35	Sequence 35, Appl	209	14	10.6	2992	3	US-09-362-123A-3	Sequence 3, Appl
137	14	10.6	1611	4	US-09-039-762A-35	Sequence 35, Appl	210	14	10.6	3016	2	US-08-344-155C-97	Sequence 97, Appl
138	14	10.6	1611	4	US-09-042-492D-35	Sequence 35, Appl	211	14	10.6	3016	2	US-09-009-490A-86	Sequence 86, Appl
139	14	10.6	1611	4	US-09-042-492D-35	Sequence 35, Appl	212	14	10.6	3024	6	5284931-1	Patent No. 5284931
140	14	10.6	1644	4	US-08-913-612A-35	Sequence 23, Appl	213	14	10.6	3066	1	US-08-142-439A-1	Sequence 1, Appl
141	14	10.6	1644	4	US-09-255-829-23	Sequence 23, Appl	214	14	10.6	3066	1	US-08-869-477-1	Sequence 1, Appl
142	14	10.6	1656	3	US-08-318-039A-2	Sequence 2, Appl	215	14	10.6	3249	1	US-08-106-493A-1	Sequence 1, Appl
143	14	10.6	1656	3	US-08-318-039B-16	Sequence 16, Appl	216	14	10.6	3249	1	US-08-429-264-1	Sequence 1, Appl
144	14	10.6	1656	3	US-08-227-496C-20	Sequence 20, Appl	217	14	10.6	3486	5	PCT-US91-09784-3	Sequence 3, Appl
145	14	10.6	1656	3	US-08-435-568A-2	Sequence 2, Appl	218	14	10.6	3507	2	US-08-775-009-36	Sequence 36, Appl
146	14	10.6	1664	4	US-09-045-583-4	Sequence 4, Appl	219	14	10.6	3509	4	US-09-255-829-19	Sequence 19, Appl
147	14	10.6	1664	4	US-09-534-185-4	Sequence 4, Appl	220	14	10.6	3660	4	US-09-517-467B-3	Sequence 3, Appl
148	14	10.6	1788	2	US-08-559-505-1	Sequence 1, Appl	221	14	10.6	3747	1	US-08-044-618-5	Sequence 5, Appl
149	14	10.6	1788	2	US-08-749-907-1	Sequence 1, Appl	222	14	10.6	3792	4	US-08-860-886-1	Sequence 1, Appl
150	14	10.6	1788	4	US-08-265-428-3	Sequence 3, Appl	223	14	10.6	3823	4	US-09-453-702B-132	Sequence 132, App
151	14	10.6	1788	5	PCT-US95-07721-3	Sequence 3, Appl	224	14	10.6	4104	1	US-07-998-003A-94	Sequence 94, Appl
152	14	10.6	1811	1	US-08-848-252-1	Sequence 1, Appl	225	14	10.6	4104	1	US-08-453-274B-94	Sequence 94, Appl
153	14	10.6	1825	2	US-08-890-980-3	Sequence 3, Appl	226	14	10.6	4104	1	US-08-453-699A-94	Sequence 94, Appl
154	14	10.6	1825	3	US-08-890-979-3	Sequence 3, Appl	227	14	10.6	4104	1	US-08-268-161A-94	Sequence 94, Appl
155	14	10.6	1825	3	US-09-032-894-3	Sequence 3, Appl	228	14	10.6	4104	2	US-08-453-702A-94	Sequence 94, Appl
156	14	10.6	1825	4	US-09-031-626-3	Sequence 3, Appl	229	14	10.6	4104	4	US-09-099-639-94	Sequence 94, Appl
157	14	10.6	1846	1	US-08-483-389-117	Sequence 117, App	230	14	10.6	4104	5	PCT-US93-12588-94	Sequence 94, Appl
158	14	10.6	1899	4	US-08-965-762-10	Sequence 10, Appl	231	14	10.6	4104	5	PCT-US95-08071-94	Sequence 94, Appl
159	14	10.6	1899	4	US-09-911-927-10	Sequence 10, Appl	232	14	10.6	4342	4	US-09-338-907-107	Sequence 107, App
160	14	10.6	1899	4	US-09-911-927-12	Sequence 12, Appl	233	14	10.6	4342	4	US-09-218-207-107	Sequence 107, App
161	14	10.6	1899	4	US-09-911-882-10	Sequence 10, Appl	234	14	10.6	4582	4	US-09-338-907-118	Sequence 118, App
162	14	10.6	1899	4	US-09-911-882-12	Sequence 12, Appl	235	14	10.6	4582	4	US-09-218-207-118	Sequence 118, App
163	14	10.6	1944	6	5260223-2	Patent No. 5260223	236	14	10.6	4686	4	US-09-338-907-117	Sequence 117, App
164	14	10.6	2000	4	US-09-605-785-374	Sequence 374, App	237	14	10.6	4686	4	US-09-218-207-117	Sequence 117, App
165	14	10.6	2000	4	US-09-439-313-374	Sequence 374, App	238	14	10.6	4823	2	US-08-457-254-5	Sequence 5, Appl
166	14	10.6	2000	4	US-09-352-616A-374	Sequence 374, App	239	14	10.6	4823	3	US-08-484-257-20	Sequence 20, Appl
167	14	10.6	2040	4	US-09-605-785-375	Sequence 375, App	240	14	10.6	4823	3	US-08-999-927-5	Sequence 5, Appl
168	14	10.6	2040	4	US-09-439-313-375	Sequence 375, App	241	14	10.6	4823	4	US-08-461-819-5	Sequence 5, Appl
169	14	10.6	2040	4	US-09-352-616A-375	Sequence 375, App	242	14	10.6	4823	5	PCT-US94-08806-28	Sequence 28, Appl
170	14	10.6	2043	3	US-08-227-496C-14	Sequence 14, Appl	243	14	10.6	4823	5	PCT-US95-01829-5	Sequence 5, Appl
171	14	10.6	2074	2	US-09-018-576-2	Sequence 2, Appl	244	14	10.6	4823	5	PCT-US95-16626-5	Sequence 5, Appl
172	14	10.6	2074	2	US-09-018-576-11	Sequence 11, Appl	245	14	10.6	4853	1	US-08-832-883-1	Sequence 1, Appl
173	14	10.6	2074	3	US-09-248-137-2	Sequence 2, Appl	246	14	10.6	4853	2	US-08-832-877-1	Sequence 1, Appl

c 247 14 10.6 4875 4 US-09-338-907-114 Sequence 114, App 320
c 248 14 10.6 4875 4 US-09-218-207-114 Sequence 114, App 321
c 249 14 10.6 4958 4 US-09-338-907-116 Sequence 116, App 322
c 250 14 10.6 4958 4 US-09-218-207-116 Sequence 116, App 323
c 251 14 10.6 4986 4 US-09-338-907-121 Sequence 121, App 324
c 252 14 10.6 4986 4 US-09-218-207-121 Sequence 121, App 325
c 253 14 10.6 5020 4 US-09-338-907-120 Sequence 120, App c 326
c 254 14 10.6 5020 4 US-09-218-207-120 Sequence 120, App 327
c 255 14 10.6 5044 4 US-09-338-907-115 Sequence 115, App 328
c 256 14 10.6 5044 4 US-09-218-207-115 Sequence 115, App 329
c 257 14 10.6 5057 4 US-09-338-907-123 Sequence 123, App c 330
c 258 14 10.6 5057 4 US-09-218-207-123 Sequence 123, App 331
c 259 14 10.6 5100 4 US-09-338-907-122 Sequence 122, App 332
c 260 14 10.6 5100 4 US-09-218-207-122 Sequence 122, App 333
c 261 14 10.6 5148 4 US-09-338-907-112 Sequence 112, App 334
c 262 14 10.6 5148 4 US-09-218-207-112 Sequence 112, App 335
c 263 14 10.6 5227 2 US-08-996-306-3 Sequence 3, Appli c 336
c 264 14 10.6 5234 4 US-09-338-907-113 Sequence 113, App 337
c 265 14 10.6 5234 4 US-09-218-207-113 Sequence 113, App 338
c 266 14 10.6 5245 4 US-09-338-907-3 Sequence 3, Appli 339
c 267 14 10.6 5245 4 US-09-218-207-3 Sequence 3, Appli 340
c 268 14 10.6 5250 4 US-09-338-907-69 Sequence 69, Appl c 341
c 269 14 10.6 5250 4 US-09-218-207-69 Sequence 69, Appl c 342
c 270 14 10.6 5290 4 US-09-338-907-119 Sequence 119, App 343
c 271 14 10.6 5290 4 US-09-218-207-119 Sequence 119, App 344
c 272 14 10.6 5326 4 US-09-338-907-124 Sequence 124, App 345
c 273 14 10.6 5326 4 US-09-218-207-124 Sequence 124, App 346
c 274 14 10.6 5345 1 US-08-044-618-7 Sequence 7, Appli c 347
c 275 14 10.6 5345 1 US-08-961-527-105 Sequence 105, App 348
c 276 14 10.6 7568 2 US-08-694-869-2 Sequence 2, Appli c 349
c 277 14 10.6 7568 3 US-09-349-546-2 Sequence 2, Appli c 350
c 278 14 10.6 8201 1 US-08-253-155A-9 Sequence 9, Appli c 351
c 279 14 10.6 8396 4 US-09-328-174A-1 Sequence 1, Appli c 352
c 280 14 10.6 8409 4 US-09-167-681-37 Sequence 37, Appl c 353
c 281 14 10.6 8697 4 US-08-961-527-123 Sequence 123, App 354
c 282 14 10.6 8697 4 US-09-385-222A-3 Sequence 3, Appli c 355
c 283 14 10.6 9836 4 US-08-972-927-2 Sequence 2, Appli c 356
c 284 14 10.6 10061 4 US-09-221-017B-672 Sequence 672, App 357
c 285 14 10.6 11236 1 US-07-853-913-1 Sequence 1, Appli c 358
c 286 14 10.6 11725 2 US-08-756-506-1 Sequence 1, Appli c 359
c 287 14 10.6 22306 4 US-09-453-702B-251 Sequence 251, App c 360
c 288 14 10.6 43676 3 US-09-356-952-12 Sequence 12, Appl 361
c 289 14 10.6 43676 3 US-09-356-952-12 Sequence 12, Appl 362
c 290 14 10.6 46819 4 US-09-453-702B-72 Sequence 72, Appl 363
c 291 14 10.6 50000 4 US-09-146-053-4 Sequence 4, Appli 364
c 292 14 10.6 56516 2 US-08-996-306-1 Sequence 1, Appli 365
c 293 14 10.6 56516 4 US-09-338-907-1 Sequence 1, Appli 366
c 294 14 10.6 56516 4 US-09-218-207-1 Sequence 1, Appli 367
c 295 14 10.6 56520 4 US-09-338-907-179 Sequence 179, App 368
c 296 14 10.6 56520 4 US-09-218-207-179 Sequence 179, App 369
c 297 14 10.6 70000 4 US-09-851-896-3 Sequence 3, Appli c 370
c 298 14 10.6 72604 4 US-09-268-992-7 Sequence 7, Appli c 371
c 299 14 10.6 72604 4 US-09-657-474-7 Sequence 7, Appli c 372
c 300 13 9.8 15 2 US-08-292-620A-110 Sequence 110, App 373
c 301 13 9.8 15 3 US-09-071-845-110 Sequence 110, App 374
c 302 13 9.8 20 1 US-08-547-182-10 Sequence 10, Appl 375
c 303 13 9.8 27 1 US-08-758-306-50 Sequence 50, Appl 376
c 304 13 9.8 27 1 US-08-758-306-50 Sequence 50, Appl 377
c 305 13 9.8 27 1 US-08-758-306-50 Sequence 50, Appl 378
c 306 13 9.8 27 1 US-08-758-306-1036 Sequence 1036, Ap 379
c 307 13 9.8 27 3 US-08-722-719-44 Sequence 44, Appl 380
c 308 13 9.8 27 4 US-08-584-040-772 Sequence 772, App 381
c 309 13 9.8 27 4 US-09-334-951-44 Sequence 44, Appl 382
c 310 13 9.8 30 2 US-08-821-782-32 Sequence 32, Appl 383
c 311 13 9.8 36 1 US-09-292-435A-32 Sequence 32, Appl 384
c 312 13 9.8 36 1 US-07-960-510-10 Sequence 10, Appl 385
c 313 13 9.8 36 1 US-07-960-510-14 Sequence 14, Appl 386
c 314 13 9.8 38 1 US-08-390-850-967 Sequence 967, App 387
c 315 13 9.8 38 1 US-08-435-634-967 Sequence 967, App 388
c 316 13 9.8 50 4 US-09-590-867A-37 Sequence 37, Appl 389
c 317 13 9.8 51 4 US-09-548-260-37 Sequence 37, Appl 390
c 318 13 9.8 52 4 US-09-461-697-98 Sequence 98, Appl 391
c 319 13 9.8 50 2 US-08-679-645-518 Sequence 518, App 392
c 320 13 9.8 50 2 US-08-927-561-109 Sequence 109, App 393

5 PCT-US96-09388-109 Sequence 109, App
1 US-08-487-141B-109 Sequence 109, App
95 4 US-08-952-793-259 Sequence 259, App
5 PCT-US96-09455A-259 Sequence 259, App
152 4 US-09-568-816A-15 Sequence 15, Appl
159 4 US-09-461-697-94 Sequence 94, Appl
192 4 US-09-134-001C-49 Sequence 49, Appl
213 3 US-08-513-974B-37 Sequence 37, Appl
213 3 US-08-943-731-160 Sequence 160, App
225 2 US-08-604-989A-7 Sequence 7, Appl
228 2 US-08-619-542B-31 Sequence 31, Appl
239 4 US-09-439-261-29 Sequence 29, Appl
239 4 US-09-227-613-28 Sequence 28, Appl
240 3 US-08-513-974B-318 Sequence 318, App
246 4 US-09-397-787-110 Sequence 110, App
254 2 US-08-605-163-21 Sequence 21, Appl
288 2 US-08-716-942-17 Sequence 17, Appl
288 4 US-09-130-337A-17 Sequence 17, Appl
291 1 US-08-594-031-147 Sequence 147, App
295 1 US-08-594-031-128 Sequence 128, App
314 4 US-08-651-155B-148 Sequence 148, App
337 1 US-08-594-031-167 Sequence 167, App
337 1 US-08-594-031-168 Sequence 168, App
340 3 US-08-441-971-7 Sequence 7, Appl
340 4 US-08-221-653-7 Sequence 7, Appl
340 4 US-08-442-144A-7 Sequence 7, Appl
340 4 US-08-441-970-7 Sequence 7, Appl
341 4 US-08-943-731-158 Sequence 158, App
355 4 US-09-155-770-1 Sequence 1, Appl
356 1 US-08-297-633A-2 Sequence 2, Appl
356 1 US-08-485-721-10 Sequence 10, Appl
356 2 US-08-392-935-10 Sequence 10, Appl
356 5 PCT-US93-08325-2 Sequence 2, Appl
356 5 PCT-US93-08326-10 Sequence 10, Appl
386 4 US-09-347-833-7 Sequence 7, Appl
428 4 US-09-456-830-32 Sequence 32, Appl
428 4 US-09-456-830-48 Sequence 48, Appl
428 4 US-09-456-830-60 Sequence 60, Appl
428 4 US-09-002-285-32 Sequence 32, Appl
428 4 US-09-002-285-48 Sequence 48, Appl
428 4 US-09-002-285-60 Sequence 60, Appl
429 4 US-09-397-787-252 Sequence 252, App
489 4 US-09-370-838-109 Sequence 109, App
496 1 US-08-263-413-23 Sequence 23, Appl
500 1 US-08-263-413-22 Sequence 22, Appl
510 4 US-09-461-697-96 Sequence 96, Appl
517 4 US-09-404-879A-211 Sequence 211, App
519 2 US-08-503-226B-36 Sequence 36, Appl
519 3 US-08-721-458B-36 Sequence 36, Appl
536 1 US-07-996-772A-5 Sequence 5, Appl
536 1 US-08-446-822-5 Sequence 5, Appl
536 4 US-09-328-314-5 Sequence 5, Appl
536 5 PCT-US93-12586-5 Sequence 5, Appl
541 2 US-08-503-226B-35 Sequence 35, Appl
541 3 US-08-721-458B-35 Sequence 35, Appl
566 4 US-09-605-785-393 Sequence 393, App
566 4 US-09-439-313-393 Sequence 393, App
566 4 US-09-352-616A-393 Sequence 393, App
570 4 US-09-328-111-461 Sequence 461, App
600 1 US-08-602-262-4 Sequence 4, Appl
600 1 US-08-602-262-5 Sequence 5, Appl
600 3 US-09-004-716-4 Sequence 4, Appl
600 3 US-09-004-716-5 Sequence 5, Appl
600 4 US-09-461-697-92 Sequence 92, Appl
608 3 US-08-115-753-6 Sequence 6, Appl
615 1 US-08-247-946A-5 Sequence 5, Appl
615 5 PCT-US95-06420-5 Patent No. 5175384
621 6 5175384-1 Patent No. 5175384
621 6 543430-1 Patent No. 543430
660 1 US-08-233-146-4 Sequence 4, Appl
660 1 US-08-463-470-4 Sequence 4, Appl
675 1 US-07-807-043B-2 Sequence 2, Appl

393	13	9.8	675	1	US-08-299-849B-2	Sequence 2, Appl	c 466	13	9.8	1252	4	US-08-785-271-13	Sequence 13, Appl
394	13	9.8	675	2	US-08-142-368A-2	Sequence 2, Appl	467	13	9.8	1271	1	US-08-464-339A-1	Sequence 1, Appl
395	13	9.8	675	3	US-08-967-727-2	Sequence 2, Appl	468	13	9.8	1271	5	PCT-US94-14388-1	Sequence 1, Appl
396	13	9.8	675	4	US-08-037-230D-2	Sequence 2, Appl	c 469	13	9.8	1279	3	US-08-826-611-5	Sequence 5, Appl
397	13	9.8	676	1	US-08-533-669A-3	Sequence 3, Appl	470	13	9.8	1287	4	US-09-134-001C-48	Sequence 48, Appl
398	13	9.8	676	2	US-08-183-861-3	Sequence 3, Appl	471	13	9.8	1292	4	US-09-198-603C-7	Sequence 7, Appl
399	13	9.8	676	4	US-09-022-765-3	Sequence 3, Appl	472	13	9.8	1294	4	US-09-425-578-1	Sequence 1, Appl
400	13	9.8	700	4	US-09-325-932A-99	Sequence 99, Appl	473	13	9.8	1302	4	US-09-134-001C-828	Sequence 828, App
401	13	9.8	737	1	US-08-602-262-3	Sequence 1, Appl	c 474	13	9.8	1316	1	US-08-446-822-7	Sequence 7, Appl
402	13	9.8	737	1	US-08-602-262-3	Sequence 1, Appl	c 475	13	9.8	1316	1	US-09-328-314-7	Sequence 7, Appl
403	13	9.8	737	3	US-09-004-716-1	Sequence 1, Appl	c 476	13	9.8	1316	5	PCT-US93-12586-7	Sequence 7, Appl
404	13	9.8	737	3	US-09-004-716-3	Sequence 3, Appl	c 477	13	9.8	1319	2	US-08-484-993B-17	Sequence 17, Appl
405	13	9.8	740	4	US-09-221-017B-355	Sequence 355, App	c 478	13	9.8	1319	2	US-08-484-158B-17	Sequence 17, Appl
406	13	9.8	756	4	US-08-413-974-3	Sequence 3, Appl	c 479	13	9.8	1319	2	US-08-484-596A-17	Sequence 17, Appl
407	13	9.8	756	4	US-08-434-418-3	Sequence 3, Appl	c 480	13	9.8	1319	2	US-08-480-150A-17	Sequence 17, Appl
408	13	9.8	756	4	US-08-433-288-3	Sequence 3, Appl	c 481	13	9.8	1319	3	US-08-458-731-17	Sequence 17, Appl
409	13	9.8	756	4	US-08-174-739A-3	Sequence 3, Appl	c 482	13	9.8	1319	3	US-08-149-223A-17	Sequence 17, Appl
410	13	9.8	756	4	US-08-434-256-3	Sequence 3, Appl	483	13	9.8	1320	6	517197-2	Patent No. 517197
411	13	9.8	792	1	US-08-403-545-6	Sequence 6, Appl	484	13	9.8	1323	4	US-08-858-207A-50	Sequence 50, Appl
412	13	9.8	792	1	US-08-446-822-14	Sequence 14, Appl	485	13	9.8	1327	4	US-09-149-476-286	Sequence 286, App
413	13	9.8	792	4	US-08-404-381-6	Sequence 6, Appl	486	13	9.8	1330	2	US-08-933-750C-80	Sequence 80, Appl
414	13	9.8	792	4	US-09-328-314-14	Sequence 14, Appl	487	13	9.8	1330	3	US-09-234-613-80	Sequence 80, Appl
415	13	9.8	792	5	PCT-US93-12586-14	Sequence 14, Appl	c 488	13	9.8	1334	2	US-08-481-658B-44	Sequence 44, Appl
416	13	9.8	801	2	US-08-903-624-5	Sequence 5, Appl	c 489	13	9.8	1334	2	US-08-477-504A-44	Sequence 44, Appl
417	13	9.8	813	4	US-08-961-527-37	Sequence 37, App	c 490	13	9.8	1334	2	US-08-486-756A-44	Sequence 44, Appl
418	13	9.8	844	4	US-09-889-595-2	Sequence 2, Appl	c 491	13	9.8	1334	3	US-08-485-862B-44	Sequence 44, Appl
419	13	9.8	879	4	US-09-221-017B-1041	Sequence 1041, Ap	c 492	13	9.8	1334	3	US-08-787-739-44	Sequence 44, Appl
420	13	9.8	883	4	US-09-171-209-74	Sequence 74, Appl	c 493	13	9.8	1334	3	US-08-487-077A-44	Sequence 44, Appl
421	13	9.8	888	3	US-09-188-930-13	Sequence 13, Appl	c 494	13	9.8	1334	3	US-08-485-863A-44	Sequence 44, Appl
422	13	9.8	909	4	US-08-928-941D-19	Sequence 19, Appl	c 495	13	9.8	1334	4	US-08-485-045D-44	Sequence 44, Appl
423	13	9.8	909	4	US-09-280-590A-19	Sequence 19, Appl	c 496	13	9.8	1334	4	US-09-178-115-44	Sequence 44, Appl
424	13	9.8	921	4	US-09-091-952A-7	Sequence 7, Appl	c 497	13	9.8	1334	4	US-09-177-776-44	Sequence 44, Appl
425	13	9.8	952	1	US-09-134-001C-2430	Sequence 2430, Ap	c 498	13	9.8	1345	1	US-08-525-654A-134	Sequence 134, App
426	13	9.8	952	1	US-08-181-271A-38	Sequence 38, Appl	499	13	9.8	1354	1	US-09-149-476-207	Sequence 207, App
427	13	9.8	952	1	US-08-449-315-38	Sequence 38, Appl	500	13	9.8	1365	1	US-07-807-043B-4	Sequence 4, Appl
428	13	9.8	952	1	US-08-444-803-38	Sequence 38, Appl	501	13	9.8	1365	1	US-08-299-849B-4	Sequence 4, Appl
429	13	9.8	952	1	US-08-449-043-38	Sequence 38, Appl	502	13	9.8	1365	2	US-08-142-368A-4	Sequence 4, Appl
430	13	9.8	952	1	US-08-456-265A-38	Sequence 38, Appl	503	13	9.8	1365	3	US-08-967-727-4	Sequence 4, Appl
431	13	9.8	952	1	US-08-455-416-38	Sequence 38, Appl	504	13	9.8	1365	4	US-08-037-230D-4	Sequence 4, Appl
432	13	9.8	952	1	US-08-455-244-38	Sequence 38, Appl	505	13	9.8	1384	4	US-08-729-594A-33	Sequence 33, Appl
433	13	9.8	952	1	US-08-454-876-38	Sequence 38, Appl	506	13	9.8	1384	4	US-08-937-993-33	Sequence 33, Appl
434	13	9.8	952	2	US-08-457-364-38	Sequence 38, Appl	507	13	9.8	1386	4	US-09-149-476-147	Sequence 147, App
435	13	9.8	952	2	US-08-456-262-38	Sequence 38, Appl	c 508	13	9.8	1398	2	US-08-604-989A-9	Sequence 9, Appl
436	13	9.8	952	2	US-08-456-240-38	Sequence 38, Appl	c 509	13	9.8	1401	1	US-07-843-945A-3	Sequence 3, Appl
437	13	9.8	952	2	US-08-455-736-38	Sequence 38, Appl	510	13	9.8	1401	2	US-08-218-978-3	Sequence 3, Appl
438	13	9.8	952	2	US-08-971-217-38	Sequence 38, Appl	c 511	13	9.8	1416	1	US-08-535-237-1	Sequence 1, Appl
439	13	9.8	952	4	US-09-350-600-38	Sequence 38, Appl	512	13	9.8	1419	2	US-08-619-542B-29	Sequence 29, Appl
440	13	9.8	1001	4	US-09-641-638-314	Sequence 314, App	513	13	9.8	1443	1	US-08-076-089-1	Sequence 1, Appl
441	13	9.8	1050	1	US-08-599-252-81	Sequence 81, Appl	514	13	9.8	1443	2	US-08-707-200-1	Sequence 1, Appl
442	13	9.8	1050	1	US-08-436-074-54	Sequence 54, Appl	515	13	9.8	1443	4	US-08-996-565-1	Sequence 1, Appl
443	13	9.8	1050	5	PCT-US96-06352-81	Sequence 81, Appl	516	13	9.8	1443	5	PCT-US93-05643-1	Sequence 1, Appl
444	13	9.8	1050	5	PCT-US96-06383-81	Sequence 81, Appl	517	13	9.8	1446	3	US-08-787-091-1	Sequence 1, Appl
445	13	9.8	1092	4	US-09-461-697-80	Sequence 80, Appl	518	13	9.8	1474	4	US-08-858-207A-144	Sequence 144, App
446	13	9.8	1123	1	US-07-971-096-3	Sequence 3, Appl	519	13	9.8	1486	4	US-09-149-476-23	Sequence 23, Appl
447	13	9.8	1123	1	US-08-175-096-3	Sequence 3, Appl	c 520	13	9.8	1508	3	US-08-867-381A-48	Sequence 48, Appl
448	13	9.8	1161	1	US-08-086-439C-1	Sequence 1, Appl	c 521	13	9.8	1508	4	US-09-521-144-48	Sequence 48, Appl
449	13	9.8	1161	1	US-08-434-877-1	Sequence 1, Appl	c 522	13	9.8	1521	2	US-08-604-989A-10	Sequence 10, Appl
450	13	9.8	1168	4	US-09-149-476-79	Sequence 79, Appl	c 523	13	9.8	1535	1	US-08-480-510-1	Sequence 1, Appl
451	13	9.8	1180	3	US-08-897-236-10	Sequence 10, Appl	c 524	13	9.8	1535	5	PCT-US94-01780-1	Sequence 1, Appl
452	13	9.8	1180	4	US-09-167-874-10	Sequence 10, Appl	c 525	13	9.8	1537	4	US-09-149-476-311	Sequence 311, App
453	13	9.8	1200	1	US-08-592-126-125	Sequence 125, App	c 526	13	9.8	1547	4	US-09-356-818A-1	Sequence 1, Appl
454	13	9.8	1206	2	US-08-666-367B-2	Sequence 2, Appl	527	13	9.8	1554	1	US-08-469-486-1	Sequence 1, Appl
455	13	9.8	1206	4	US-09-143-438-2	Sequence 2, Appl	528	13	9.8	1554	2	US-08-469-658-1	Sequence 1, Appl
456	13	9.8	1224	1	US-09-574-942-1	Sequence 1, Appl	c 529	13	9.8	1601	3	US-08-817-188-3	Sequence 3, Appl
457	13	9.8	1225	1	US-08-547-182-1	Sequence 1, Appl	530	13	9.8	1603	4	US-08-065-844A-1	Sequence 1, Appl
458	13	9.8	1239	4	US-09-461-697-76	Sequence 76, Appl	c 531	13	9.8	1605	4	US-09-149-476-187	Sequence 187, App
459	13	9.8	1246	4	US-09-651-656-30	Sequence 30, Appl	c 532	13	9.8	1622	1	US-07-996-772A-3	Sequence 3, Appl
460	13	9.8	1246	4	US-09-650-855-30	Sequence 30, Appl	c 533	13	9.8	1622	1	US-08-446-822-3	Sequence 3, Appl
461	13	9.8	1251	4	US-09-242-095-1	Sequence 1, Appl	c 534	13	9.8	1622	4	US-09-328-314-3	Sequence 3, Appl
462	13	9.8	1252	2	US-08-588-983-43	Sequence 43, Appl	c 535	13	9.8	1622	5	PCT-US93-12586-3	Sequence 3, Appl
463	13	9.8	1252	3	US-08-588-976-43	Sequence 43, Appl	c 536	13	9.8	1642	1	US-07-996-772A-1	Sequence 1, Appl
464	13	9.8	1252	3	US-08-589-028-13	Sequence 13, Appl	c 537	13	9.8	1642	1	US-08-446-822-1	Sequence 1, Appl
465	13	9.8	1252	3	US-08-784-582-13	Sequence 13, Appl	c 538	13	9.8	1642	4	US-09-328-314-1	Sequence 1, Appl

c 539	13	9.8	1642	5	PCT-US93-12586-1	Sequence 1, Appli	612	13	9.8	2073	4	US-09-173-300-3	Sequence 3, Appli
c 540	13	9.8	1645	2	US-08-461-812-1	Sequence 1, Appli	613	13	9.8	2089	4	US-09-155-770-6	Sequence 6, Appli
c 541	13	9.8	1673	4	US-09-426-568A-1	Sequence 1, Appli	c 614	13	9.8	2098	2	US-08-937-931-5	Sequence 5, Appli
c 542	13	9.8	1679	4	US-09-271-437-3	Sequence 3, Appli	c 615	13	9.8	2098	4	US-09-285-502-5	Sequence 5, Appli
c 543	13	9.8	1683	2	US-08-838-543-1	Sequence 1, Appli	c 616	13	9.8	2098	4	US-09-709-126-5	Sequence 5, Appli
c 544	13	9.8	1704	4	US-09-608-917A-3	Sequence 3, Appli	c 617	13	9.8	2098	4	US-09-871-385A-5	Sequence 5, Appli
c 545	13	9.8	1710	4	US-09-741-154-1	Sequence 1, Appli	618	13	9.8	2105	4	US-09-088-425-3	Sequence 3, Appli
c 546	13	9.8	1730	2	US-08-687-080-114	Sequence 11A, App	619	13	9.8	2114	4	US-09-568-816A-1	Sequence 3, Appli
c 547	13	9.8	1734	2	US-08-948-569A-7	Sequence 7, Appli	c 620	13	9.8	2127	2	US-08-832-883-54	Sequence 54, Appl
c 548	13	9.8	1734	2	US-08-663-808-1	Sequence 1, Appli	c 621	13	9.8	2127	2	US-08-832-877-54	Sequence 54, Appl
c 549	13	9.8	1734	2	US-09-188-469-7	Sequence 7, Appli	c 622	13	9.8	2148	4	US-09-620-412C-320	Sequence 320, App
c 550	13	9.8	1734	3	US-09-332-740-1	Sequence 1, Appli	c 623	13	9.8	2155	2	US-08-031-538-3	Sequence 3, Appli
c 551	13	9.8	1734	3	US-09-188-496-1	Sequence 1, Appli	624	13	9.8	2156	2	US-08-899-514-1	Sequence 1, Appli
c 552	13	9.8	1734	4	US-09-397-238A-7	Sequence 7, Appli	625	13	9.8	2172	4	US-08-030-410-2	Sequence 2, Appli
c 553	13	9.8	1734	4	US-09-368-282-1	Sequence 1, Appli	626	13	9.8	2175	4	US-09-357-251-9	Sequence 9, Appli
c 554	13	9.8	1734	4	US-09-566-708A-1	Sequence 1, Appli	627	13	9.8	2184	1	US-07-815-333A-1	Sequence 1, Appli
c 555	13	9.8	1762	1	US-08-105-483-300	Sequence 300, App	628	13	9.8	2220	4	US-08-997-251-1	Sequence 1, Appli
c 556	13	9.8	1762	1	US-08-709-209-300	Sequence 300, App	629	13	9.8	2244	4	US-09-354-151-1	Sequence 1, Appli
c 557	13	9.8	1762	1	US-08-458-101-300	Sequence 300, App	630	13	9.8	2264	2	US-08-149-097D-26	Sequence 26, Appl
c 558	13	9.8	1786	1	US-07-920-430-19	Sequence 19, Appl	631	13	9.8	2272	3	US-08-147-592A-3	Sequence 3, Appli
c 559	13	9.8	1786	1	US-08-066-299-9	Sequence 9, Appli	632	13	9.8	2272	4	US-08-292-694A-3	Sequence 3, Appli
c 560	13	9.8	1786	1	US-08-265-047-1	Sequence 1, Appli	633	13	9.8	2272	1	US-08-369-796-7	Sequence 7, Appli
c 561	13	9.8	1786	1	US-08-251-464-19	Sequence 19, Appl	634	13	9.8	2277	2	US-08-852-091-7	Sequence 7, Appli
c 562	13	9.8	1786	4	US-08-926-522-20	Sequence 20, Appl	635	13	9.8	2277	2	US-08-820-754-7	Sequence 7, Appli
c 563	13	9.8	1786	5	PCT-US92-0136A-19	Sequence 19, Appl	636	13	9.8	2277	3	US-08-956-652-7	Sequence 7, Appli
c 564	13	9.8	1810	1	US-07-755-573C-7	Sequence 7, Appli	637	13	9.8	2277	3	US-08-956-869-7	Sequence 7, Appli
c 565	13	9.8	1820	4	US-09-530-212A-1	Sequence 1, Appli	638	13	9.8	2277	3	US-08-948-547-7	Sequence 7, Appli
c 566	13	9.8	1825	4	US-09-461-697-75	Sequence 75, Appl	639	13	9.8	2277	4	US-08-956-653A-7	Sequence 7, Appli
c 567	13	9.8	1829	2	US-08-411-859-1	Sequence 1, Appli	640	13	9.8	2277	4	US-09-347-878-53	Sequence 53, Appl
c 568	13	9.8	1829	4	US-08-387-707-7	Sequence 7, Appli	641	13	9.8	2277	5	PCT-US95-17025-7	Sequence 7, Appli
c 569	13	9.8	1829	4	US-08-405-271A-7	Sequence 7, Appli	642	13	9.8	2278	4	US-09-442-055-1	Sequence 1, Appli
c 570	13	9.8	1838	3	US-08-948-564-1	Sequence 1, Appli	643	13	9.8	2288	1	US-08-290-937B-4	Sequence 4, Appli
c 571	13	9.8	1839	3	US-08-152-019A-44	Sequence 44, Appl	644	13	9.8	2289	1	US-07-838-410-2	Sequence 2, Appli
c 572	13	9.8	1839	3	US-08-482-677-1	Sequence 1, Appli	c 645	13	9.8	2339	4	US-09-268-140-11	Sequence 11, Appl
c 573	13	9.8	1851	4	US-09-149-476-62	Sequence 62, Appl	646	13	9.8	2348	4	US-09-393-554-10	Sequence 10, Appl
c 574	13	9.8	1858	2	US-08-909-965C-11	Sequence 11, Appl	647	13	9.8	2352	4	US-08-997-251-3	Sequence 3, Appli
c 575	13	9.8	1886	1	US-07-980-526-1	Sequence 1, Appli	648	13	9.8	2408	2	US-08-870-827-5	Sequence 5, Appli
c 576	13	9.8	1890	4	US-09-134-001C-1557	Sequence 1557, Ap	649	13	9.8	2408	2	US-08-317-179-5	Sequence 1, Appli
c 577	13	9.8	1900	4	US-09-468-265-2	Sequence 2, Appli	c 650	13	9.8	2410	2	US-08-780-835B-1	Sequence 1, Appli
c 578	13	9.8	1909	4	US-09-431-099-3	Sequence 3, Appli	c 651	13	9.8	2410	4	US-09-303-268-1	Sequence 1, Appli
c 579	13	9.8	1910	4	US-08-247-902A-1	Sequence 1, Appli	c 652	13	9.8	2410	4	US-09-116-049-1	Sequence 1, Appli
c 580	13	9.8	1910	5	PCT-US93-10541-1	Sequence 1, Appli	c 653	13	9.8	2412	4	US-09-072-596-345	Sequence 345, App
c 581	13	9.8	1915	4	US-08-369-046-1	Sequence 1, Appli	654	13	9.8	2424	1	US-08-821-119-16	Sequence 16, Appl
c 582	13	9.8	1925	3	US-08-669-378-1	Sequence 1, Appli	655	13	9.8	2424	4	US-09-149-476-228	Sequence 228, App
c 583	13	9.8	1925	3	US-08-669-378-3	Sequence 3, Appli	656	13	9.8	2430	1	US-08-062-368-3	Sequence 3, Appli
c 584	13	9.8	1925	3	US-08-669-378-5	Sequence 5, Appli	657	13	9.8	2430	1	US-08-233-146-8	Sequence 8, Appli
c 585	13	9.8	1925	3	US-08-669-378-7	Sequence 7, Appli	658	13	9.8	2456	1	US-08-463-470-8	Sequence 8, Appli
c 586	13	9.8	1925	3	US-08-669-378-9	Sequence 9, Appli	c 659	13	9.8	2469	1	US-08-858-207A-51	Sequence 51, Appl
c 587	13	9.8	1925	3	US-08-669-378-11	Sequence 11, Appl	c 660	13	9.8	2483	1	US-08-464-340A-3	Sequence 3, Appli
c 588	13	9.8	1932	4	US-09-518-386B-2	Sequence 2, Appli	c 661	13	9.8	2483	5	PCT-US94-08449A-3	Sequence 3, Appli
c 589	13	9.8	1942	2	US-08-604-989A-11	Sequence 11, Appl	662	13	9.8	2505	4	US-09-268-140-1	Sequence 7, Appli
c 590	13	9.8	1956	3	US-08-996-441B-57	Sequence 57, Appl	c 663	13	9.8	2517	4	US-09-534-407-5	Sequence 5, Appli
c 591	13	9.8	1956	3	US-08-993-722A-57	Sequence 57, Appl	664	13	9.8	2517	4	US-09-534-407-5	Sequence 5, Appli
c 592	13	9.8	1956	3	US-08-993-170A-57	Sequence 57, Appl	665	13	9.8	2526	1	US-07-912-952-1	Sequence 1, Appli
c 593	13	9.8	1956	3	US-08-996-441B-57	Sequence 57, Appl	666	13	9.8	2556	1	US-08-208-887A-48	Sequence 48, Appl
c 594	13	9.8	1959	3	US-08-996-441B-1	Sequence 1, Appli	667	13	9.8	2556	3	US-08-699-103B-9	Sequence 9, Appli
c 595	13	9.8	1959	3	US-08-993-722A-1	Sequence 1, Appli	668	13	9.8	2556	4	US-09-229-059-9	Sequence 9, Appli
c 596	13	9.8	1959	3	US-08-993-775B-1	Sequence 1, Appli	669	13	9.8	2556	4	US-09-280-598-17	Sequence 17, Appl
c 597	13	9.8	1959	3	US-09-072-596-349	Sequence 349, App	670	13	9.8	2625	6	5175384-10	Patent No. 5175384
c 598	13	9.8	1974	4	US-09-291-541-9	Sequence 9, Appli	671	13	9.8	2634	3	US-08-949-386-26	Sequence 26, Appl
c 599	13	9.8	1987	2	US-08-876-882-1	Sequence 1, Appli	672	13	9.8	2634	3	US-08-450-562-26	Sequence 26, Appl
c 600	13	9.8	1987	4	US-09-315-928-1	Sequence 1, Appli	673	13	9.8	2634	4	US-08-984-709A-26	Sequence 26, Appl
c 601	13	9.8	1987	4	US-08-818-111-152	Sequence 152, App	674	13	9.8	2634	4	US-08-450-272-26	Sequence 26, Appl
c 602	13	9.8	1993	4	US-08-818-111-147	Sequence 147, App	675	13	9.8	2644	2	US-07-728-215-26	Sequence 26, Appl
c 603	13	9.8	1993	4	US-09-056-556-152	Sequence 152, App	676	13	9.8	2644	4	US-08-938-085A-26	Sequence 26, Appl
c 604	13	9.8	1993	4	US-09-056-556-152	Sequence 152, App	c 677	13	9.8	2671	2	US-08-666-367B-1	Sequence 1, Appli
c 605	13	9.8	1993	4	US-09-056-556-152	Sequence 152, App	c 678	13	9.8	2671	4	US-09-143-438-1	Sequence 1, Appli
c 606	13	9.8	1993	4	US-09-072-596-147	Sequence 147, App	679	13	9.8	2704	6	5434340-6	Patent No. 5434340
c 607	13	9.8	1993	4	US-09-072-596-149	Sequence 149, App	680	13	9.8	2712	3	US-08-949-386-38	Sequence 38, Appl
c 608	13	9.8	2000	4	US-08-426-509A-1	Sequence 1, Appli	681	13	9.8	2712	3	US-08-450-562-38	Sequence 38, Appl
c 609	13	9.8	2000	5	PCT-US95-05008-1	Sequence 1, Appli	682	13	9.8	2712	4	US-08-984-709A-38	Sequence 38, Appl
c 610	13	9.8	2032	4	US-09-045-284A-1	Sequence 1, Appli	683	13	9.8	2712	4	US-08-450-272-38	Sequence 38, Appl
c 611	13	9.8	2032	4	US-09-190-911-2	Sequence 2, Appli	684	13	9.8	2731	3	US-08-699-103B-11	Sequence 11, Appl

685	13	9.8	2731	4	US-09-229-059-11	Sequence 11, Appl	758	13	9.8	3546	3	US-08-459-595A-14	Sequence 14, Appl
686	13	9.8	2786	1	US-08-224-391-53	Sequence 53, Appl	759	13	9.8	3546	3	US-08-459-504B-10	Sequence 10, Appl
687	13	9.8	2786	1	US-08-484-304-53	Sequence 53, Appl	760	13	9.8	3546	3	US-08-459-504B-12	Sequence 12, Appl
688	13	9.8	2817	4	US-09-431-099-1	Sequence 1, Appl	761	13	9.8	3546	3	US-08-459-504B-14	Sequence 14, Appl
689	13	9.8	2844	4	US-09-257-703-2	Sequence 2, Appl	762	13	9.8	3546	3	US-08-459-444-10	Sequence 10, Appl
690	13	9.8	2897	2	US-08-927-394-1	Sequence 1, Appl	763	13	9.8	3546	3	US-08-459-444-12	Sequence 12, Appl
691	13	9.8	2903	4	US-08-928-941D-2	Sequence 2, Appl	764	13	9.8	3546	3	US-08-459-444-14	Sequence 14, Appl
692	13	9.8	2903	4	US-08-928-941D-3	Sequence 3, Appl	765	13	9.8	3546	3	US-09-547-422-10	Sequence 10, Appl
693	13	9.8	2903	4	US-09-280-590A-2	Sequence 2, Appl	766	13	9.8	3546	4	US-09-547-422-12	Sequence 12, Appl
694	13	9.8	2903	4	US-09-280-590A-3	Sequence 3, Appl	767	13	9.8	3546	4	US-09-547-422-14	Sequence 14, Appl
695	13	9.8	2938	4	US-09-072-917A-8	Sequence 8, Appl	768	13	9.8	3572	4	US-09-072-936-337	Sequence 337, App
696	13	9.8	2946	4	US-09-175-928-3	Sequence 3, Appl	769	13	9.8	3627	1	US-08-104-072B-6	Sequence 6, Appl
697	13	9.8	2952	4	US-09-134-001C-977	Sequence 977, App	770	13	9.8	3627	1	US-08-351-413-7	Sequence 7, Appl
698	13	9.8	2970	3	US-08-949-386-37	Sequence 37, Appl	771	13	9.8	3627	2	US-09-025-583-7	Sequence 7, Appl
699	13	9.8	2970	3	US-08-450-562-37	Sequence 37, Appl	772	13	9.8	3668	2	US-08-849-345-1	Sequence 1, Appl
700	13	9.8	2970	4	US-08-984-709A-37	Sequence 37, Appl	773	13	9.8	3931	2	US-08-144-121-1	Sequence 1, Appl
701	13	9.8	2970	4	US-08-450-273-37	Sequence 37, Appl	774	13	9.8	3931	2	US-08-735-893-1	Sequence 1, Appl
702	13	9.8	2973	4	US-09-133-962A-15	Sequence 15, Appl	775	13	9.8	3969	4	US-09-518-386B-4	Sequence 4, Appl
703	13	9.8	3120	1	US-08-222-616-22	Sequence 22, Appl	776	13	9.8	4016	5	PCT-US95-08354A-1	Sequence 1, Appl
704	13	9.8	3120	4	US-08-446-648-22	Sequence 22, Appl	777	13	9.8	4052	1	US-08-057-167-1	Sequence 1, Appl
705	13	9.8	3120	5	PCT-US95-04228-22	Sequence 22, Appl	778	13	9.8	4052	5	PCT-US93-05412-1	Sequence 1, Appl
706	13	9.8	3142	4	US-09-221-017B-869	Sequence 869, App	779	13	9.8	4069	4	US-09-302-812-3	Sequence 3, Appl
707	13	9.8	3156	2	US-08-887-518-1	Sequence 1, Appl	780	13	9.8	4069	4	US-09-511-477-3	Sequence 3, Appl
708	13	9.8	3156	2	US-09-023-321-1	Sequence 1, Appl	781	13	9.8	4069	4	US-09-511-507-3	Sequence 3, Appl
709	13	9.8	3156	2	US-09-032-475-1	Sequence 1, Appl	782	13	9.8	4141	4	US-09-245-281-42	Sequence 42, Appl
710	13	9.8	3257	5	PCT-US91-09784-1	Sequence 1, Appl	783	13	9.8	4141	4	US-09-207-359B-42	Sequence 42, Appl
711	13	9.8	3288	4	US-09-206-942-44	Sequence 44, Appl	784	13	9.8	4533	3	US-08-863-790-27	Sequence 27, Appl
712	13	9.8	3306	4	US-09-206-942-42	Sequence 42, Appl	785	13	9.8	4533	3	US-08-296-749-27	Sequence 27, Appl
713	13	9.8	3366	4	US-09-171-461-27	Sequence 27, Appl	786	13	9.8	4637	4	US-09-221-017B-818	Sequence 818, App
714	13	9.8	3375	4	US-09-309-573-1	Sequence 1, Appl	787	13	9.8	4698	1	US-07-807-043B-5	Sequence 5, Appl
715	13	9.8	3376	4	US-09-309-573-3	Sequence 3, Appl	788	13	9.8	4698	1	US-08-299-849B-5	Sequence 5, Appl
716	13	9.8	3403	4	US-08-448-489-2	Sequence 2, Appl	789	13	9.8	4698	2	US-08-142-368A-5	Sequence 5, Appl
717	13	9.8	3454	4	US-09-082-059-1	Sequence 1, Appl	790	13	9.8	4698	3	US-08-967-727-5	Sequence 5, Appl
718	13	9.8	3468	1	US-07-951-715A-2	Sequence 2, Appl	791	13	9.8	4698	3	US-08-037-230D-5	Sequence 5, Appl
719	13	9.8	3468	1	US-07-951-715A-4	Sequence 4, Appl	792	13	9.8	4698	4	US-09-439-261-34	Sequence 34, Appl
720	13	9.8	3468	2	US-08-459-448A-2	Sequence 2, Appl	793	13	9.8	4698	4	US-09-227-613-33	Sequence 33, Appl
721	13	9.8	3468	3	US-08-459-488A-4	Sequence 4, Appl	794	13	9.8	4700	4	US-09-150-460B-9	Sequence 9, Appl
722	13	9.8	3468	3	US-08-459-595A-2	Sequence 2, Appl	795	13	9.8	4712	4	US-09-221-017B-577	Sequence 577, App
723	13	9.8	3468	3	US-08-459-595A-4	Sequence 4, Appl	796	13	9.8	4868	5	PCT-US93-11310-12	Sequence 12, Appl
724	13	9.8	3468	3	US-08-459-504B-2	Sequence 2, Appl	797	13	9.8	4868	5	PCT-US93-11310-12	Sequence 12, Appl
725	13	9.8	3468	3	US-08-459-504B-4	Sequence 4, Appl	798	13	9.8	4895	4	US-09-426-568A-3	Sequence 3, Appl
726	13	9.8	3468	3	US-08-459-444-2	Sequence 2, Appl	799	13	9.8	4900	1	US-08-245-295-5	Sequence 5, Appl
727	13	9.8	3468	3	US-08-459-444-4	Sequence 4, Appl	800	13	9.8	4900	1	US-08-481-130-5	Sequence 5, Appl
728	13	9.8	3468	3	US-09-053-549-5	Sequence 5, Appl	801	13	9.8	4900	1	US-08-656-984A-5	Sequence 5, Appl
729	13	9.8	3468	3	US-09-053-549-5	Sequence 5, Appl	802	13	9.8	4900	2	US-08-487-595-5	Sequence 5, Appl
730	13	9.8	3468	4	US-09-547-422-2	Sequence 2, Appl	803	13	9.8	4900	2	US-08-487-595-5	Sequence 5, Appl
731	13	9.8	3471	4	US-09-002-285-73	Sequence 73, Appl	804	13	9.8	4900	3	US-08-863-790-26	Sequence 26, Appl
732	13	9.8	3471	5	PCT-US93-00227-2	Sequence 2, Appl	805	13	9.8	4900	3	US-08-296-749-26	Sequence 26, Appl
733	13	9.8	3475	1	US-08-222-299-3	Sequence 3, Appl	806	13	9.8	5011	1	US-08-141-893-1	Sequence 1, Appl
734	13	9.8	3475	1	US-08-434-878-3	Sequence 3, Appl	807	13	9.8	5011	1	US-08-463-092B-1	Sequence 1, Appl
735	13	9.8	3475	5	PCT-US95-03718-3	Sequence 3, Appl	808	13	9.8	5011	1	US-08-463-092B-3	Sequence 3, Appl
736	13	9.8	3476	1	US-08-183-211-1	Sequence 1, Appl	809	13	9.8	5011	2	US-08-462-109A-1	Sequence 1, Appl
737	13	9.8	3476	5	PCT-US95-00176A-1	Sequence 1, Appl	810	13	9.8	5011	2	US-08-462-109A-3	Sequence 3, Appl
738	13	9.8	3476	5	US-07-977-451-3	Sequence 3, Appl	811	13	9.8	5011	2	US-08-460-907B-1	Sequence 1, Appl
739	13	9.8	3501	1	US-07-977-451-3	Sequence 3, Appl	812	13	9.8	5011	2	US-08-460-907B-3	Sequence 3, Appl
740	13	9.8	3501	1	US-08-253-517-3	Sequence 3, Appl	813	13	9.8	5011	3	US-08-463-179A-1	Sequence 1, Appl
741	13	9.8	3501	1	US-07-906-397A-3	Sequence 3, Appl	814	13	9.8	5011	3	US-08-463-179A-3	Sequence 3, Appl
742	13	9.8	3501	1	US-08-601-891-3	Sequence 3, Appl	815	13	9.8	5011	3	US-08-461-384B-1	Sequence 1, Appl
743	13	9.8	3501	2	PCT-US92-05401-3	Sequence 3, Appl	816	13	9.8	5011	3	US-08-461-384B-3	Sequence 3, Appl
744	13	9.8	3501	5	PCT-US92-09893-3	Sequence 3, Appl	817	13	9.8	5011	3	US-08-407-207A-1	Sequence 1, Appl
745	13	9.8	3501	5	US-08-315-468-3	Sequence 3, Appl	818	13	9.8	5058	4	US-09-889-595-1	Sequence 1, Appl
746	13	9.8	3516	2	US-08-943-087-1	Sequence 1, Appl	819	13	9.8	5077	1	US-08-485-604-8	Sequence 8, Appl
747	13	9.8	3539	4	US-09-245-248B-59	Sequence 59, Appl	820	13	9.8	5077	1	US-08-485-604-8	Sequence 8, Appl
748	13	9.8	3546	1	US-08-086-634-1	Sequence 1, Appl	821	13	9.8	5077	1	US-08-487-595-8	Sequence 8, Appl
749	13	9.8	3546	1	US-07-951-715A-10	Sequence 10, Appl	822	13	9.8	5077	1	US-08-487-595-8	Sequence 8, Appl
750	13	9.8	3546	1	US-07-951-715A-12	Sequence 12, Appl	823	13	9.8	5089	6	5177197-31	Patent No. 5177197
751	13	9.8	3546	1	US-07-951-715A-14	Sequence 14, Appl	824	13	9.8	5134	2	US-08-635-121-1	Sequence 1, Appl
752	13	9.8	3546	1	US-08-459-448A-10	Sequence 10, Appl	825	13	9.8	5152	4	US-09-690-364-10	Sequence 10, Appl
753	13	9.8	3546	2	US-08-459-448A-12	Sequence 12, Appl	826	13	9.8	5240	4	US-09-171-337A-2	Sequence 2, Appl
754	13	9.8	3546	2	US-08-459-448A-14	Sequence 14, Appl	827	13	9.8	5261	1	US-08-045-806-3	Sequence 3, Appl
755	13	9.8	3546	2	US-08-459-448A-16	Sequence 16, Appl	828	13	9.8	5261	1	US-08-366-051B-3	Sequence 3, Appl
756	13	9.8	3546	3	US-08-459-595A-10	Sequence 10, Appl	829	13	9.8	5261	1	US-08-441-139-12	Sequence 12, Appl
757	13	9.8	3546	3	US-08-459-595A-12	Sequence 12, Appl	830	13	9.8	5470	1		

831	13	9.8	5470	6	5196523-5	Patent No. 5196523	904	13	9.8	15894	1	US-08-348-891A-1	Sequence 1, Appli
c 832	13	9.8	5494	4	US-08-910-864-10	Sequence 10, Appl	905	13	9.8	15894	1	US-08-905-817-1	Sequence 1, Appli
c 833	13	9.8	5510	1	US-08-123-161A-7	Sequence 7, Appli	c 906	13	9.8	16389	4	US-09-741-154-3	Sequence 1, Appli
c 834	13	9.8	5510	1	US-08-483-278-7	Sequence 7, Appli	907	13	9.8	17341	4	US-09-415-946-1	Sequence 1, Appli
c 835	13	9.8	5578	1	US-08-081-610-2	Sequence 2, Appli	908	13	9.8	17606	4	US-08-943-731-4	Sequence 4, Appli
c 836	13	9.8	5690	2	US-08-447-464-2	Sequence 2, Appli	c 909	13	9.8	17606	4	US-08-943-731-4	Sequence 4, Appli
c 837	13	9.8	5690	2	US-08-716-679-2	Sequence 2, Appli	c 910	13	9.8	18436	4	US-08-961-527-87	Sequence 87, Appl
c 838	13	9.8	5737	1	US-08-259-264-1	Sequence 1, Appli	c 911	13	9.8	18627	4	US-08-961-527-113	Sequence 113, App
c 839	13	9.8	5838	3	US-08-578-096A-1	Sequence 1, Appli	c 912	13	9.8	19182	2	US-08-850-880-11	Sequence 11, Appl
c 840	13	9.8	5838	3	US-09-240-426-1	Sequence 1, Appli	c 913	13	9.8	19182	2	US-08-944-916-11	Sequence 11, Appl
c 841	13	9.8	5874	4	US-08-843-417-9	Sequence 9, Appli	c 914	13	9.8	19182	2	US-09-272-432A-11	Sequence 11, Appl
c 842	13	9.8	5893	1	US-08-592-126-54	Sequence 54, Appl	c 915	13	9.8	24979	3	US-08-147-777-3	Sequence 3, Appli
c 843	13	9.8	5893	2	US-08-687-080-44	Sequence 44, Appl	c 916	13	9.8	24979	3	US-08-452-872-3	Sequence 3, Appli
c 844	13	9.8	6002	1	US-08-698-551-15	Sequence 15, Appl	c 917	13	9.8	24979	5	US-08-452-872-3	Sequence 3, Appli
c 845	13	9.8	6002	2	US-08-602-228-15	Sequence 15, Appl	c 918	13	9.8	28001	4	US-09-819-993-3	Sequence 3, Appli
c 846	13	9.8	6002	2	US-08-839-032A-15	Sequence 15, Appl	c 919	13	9.8	28001	4	US-09-819-993-3	Sequence 3, Appli
c 847	13	9.8	6002	4	US-09-185-2580-15	Sequence 15, Appl	c 920	13	9.8	28720	4	US-09-341-587-7	Sequence 7, Appli
c 848	13	9.8	6122	1	US-08-403-545-1	Sequence 1, Appli	c 921	13	9.8	28720	4	US-09-341-587-7	Sequence 7, Appli
c 849	13	9.8	6122	4	US-08-404-381-1	Sequence 1, Appli	c 922	13	9.8	28882	4	US-08-961-527-140	Sequence 140, App
c 850	13	9.8	6315	2	US-08-808-793-2	Sequence 2, Appli	c 923	13	9.8	29598	4	US-09-341-587-6	Sequence 6, Appli
c 851	13	9.8	6315	3	US-08-772-512A-2	Sequence 2, Appli	c 924	13	9.8	29629	4	US-09-729-995-3	Sequence 3, Appli
c 852	13	9.8	6318	2	US-08-808-793-1	Sequence 2, Appli	c 925	13	9.8	31960	4	US-09-453-702B-11	Sequence 11, Appl
c 853	13	9.8	6318	3	US-08-772-512A-1	Sequence 1, Appli	c 926	13	9.8	36741	4	US-09-301-665-3	Sequence 3, Appli
c 854	13	9.8	6605	3	US-08-769-309A-4	Sequence 4, Appli	c 927	13	9.8	37948	4	US-09-251-645-11	Sequence 11, Appl
c 855	13	9.8	6605	3	US-08-594-570-4	Sequence 4, Appli	c 928	13	9.8	38564	4	US-09-734-673-3	Sequence 3, Appli
c 856	13	9.8	7082	4	US-09-362-831-1	Sequence 1, Appli	c 929	13	9.8	43804	4	US-09-171-461-1	Sequence 1, Appli
c 857	13	9.8	7152	4	US-09-167-681-29	Sequence 29, Appl	c 930	13	9.8	50000	4	US-09-146-053-4	Sequence 4, Appli
c 858	13	9.8	7163	4	US-08-961-527-67	Sequence 67, Appl	c 931	13	9.8	50341	2	US-08-247-901C-1	Sequence 1, Appli
c 859	13	9.8	7172	4	US-08-961-527-120	Sequence 120, App	c 932	13	9.8	50341	2	US-09-075-904-1	Sequence 1, Appli
c 860	13	9.8	7559	2	US-08-250-848-2	Sequence 2, Appli	c 933	13	9.8	51259	3	US-08-781-891-209	Sequence 209, App
c 861	13	9.8	7676	4	US-09-056-556-213	Sequence 213, App	c 934	13	9.8	52297	4	US-09-426-436-1	Sequence 1, Appli
c 862	13	9.8	7676	4	US-09-072-596-208	Sequence 208, App	c 935	13	9.8	52297	4	US-08-705-357-1	Sequence 1, Appli
c 863	13	9.8	7785	2	US-08-276-967-1	Sequence 1, Appli	c 936	13	9.8	84495	4	US-09-797-906-3	Sequence 3, Appli
c 864	13	9.8	7881	2	US-08-751-189-1	Sequence 1, Appli	c 937	13	9.8	87350	3	US-08-781-891-79	Sequence 79, Appl
c 865	13	9.8	7881	2	US-09-060-836-1	Sequence 1, Appli	c 938	13	9.8	87543	4	US-09-791-211-3	Sequence 3, Appli
c 866	13	9.8	7881	4	US-09-184-445-1	Sequence 1, Appli	c 939	13	9.8	90050	4	US-09-245-041-5	Sequence 5, Appli
c 867	13	9.8	8065	4	US-09-091-952A-6	Sequence 6, Appli	c 940	13	9.8	90050	4	US-09-245-041-5	Sequence 5, Appli
c 868	13	9.8	8703	4	US-08-961-527-177	Sequence 177, App	c 941	13	9.8	98844	4	US-09-791-211-10	Sequence 10, Appl
c 869	13	9.8	8789	1	US-08-328-254-5	Sequence 5, Appli	c 942	13	9.8	111282	4	US-09-754-250-3	Sequence 3, Appli
c 870	13	9.8	8791	5	PCT-US96-01735-5	Sequence 5, Appli	c 943	13	9.8	112132	4	US-09-741-150-3	Sequence 3, Appli
c 871	13	9.8	8854	3	US-09-053-549-1	Sequence 1, Appli	c 944	13	9.8	112132	3	US-09-741-150-3	Sequence 3, Appli
c 872	13	9.8	8916	4	US-09-579-181-11	Sequence 11, Appl	c 945	13	9.8	176373	3	US-09-128-155-17	Sequence 17, Appl
c 873	13	9.8	9354	4	US-09-579-181-10	Sequence 10, Appl	c 946	13	9.8	176373	3	US-09-128-155-17	Sequence 17, Appl
c 874	13	9.8	10136	1	US-08-353-700-2	Sequence 2, Appli	c 947	13	9.8	246240	2	US-08-724-394A-20	Sequence 20, Appl
c 875	13	9.8	10136	5	PCT-US95-16210-2	Sequence 2, Appli	c 948	13	9.8	246240	2	US-08-724-394A-20	Sequence 20, Appl
c 876	13	9.8	10357	1	US-08-961-527-191	Sequence 191, App	c 949	13	9.8	246240	2	US-08-724-394A-22	Sequence 22, Appl
c 877	13	9.8	10409	3	US-08-772-440-33	Sequence 33, Appl	c 950	13	9.8	4403765	4	US-09-103-840A-2	Sequence 22, Appl
c 878	13	9.8	10642	4	US-09-934-551-3	Sequence 3, Appli	c 951	13	9.8	4411529	3	US-09-103-840A-1	Sequence 1, Appli
c 879	13	9.8	10898	2	US-08-481-658B-5	Sequence 5, Appli	c 952	12	9.1	18	3	US-09-289-377-35	Sequence 35, Appl
c 880	13	9.8	10898	2	US-08-481-658B-5	Sequence 5, Appli	c 953	12	9.1	18	4	US-08-584-040-3067	Sequence 3067, Ap
c 881	13	9.8	10898	2	US-08-486-756A-5	Sequence 5, Appli	c 954	12	9.1	19	1	US-08-244-309-7	Sequence 7, Appli
c 882	13	9.8	10898	2	US-08-485-862B-5	Sequence 5, Appli	c 955	12	9.1	20	4	US-09-356-818A-3	Sequence 3, Appli
c 883	13	9.8	10898	3	US-08-787-739-5	Sequence 5, Appli	c 956	12	9.1	20	4	US-08-997-251-7	Sequence 7, Appli
c 884	13	9.8	10898	3	US-08-487-077A-5	Sequence 5, Appli	c 957	12	9.1	20	4	US-09-555-778-2	Sequence 2, Appli
c 885	13	9.8	10898	3	US-08-485-863A-5	Sequence 5, Appli	c 958	12	9.1	20	4	US-09-430-114-6	Sequence 6, Appli
c 886	13	9.8	10898	4	US-08-485-049D-5	Sequence 5, Appli	c 959	12	9.1	20	4	US-09-702-251-85	Sequence 85, Appl
c 887	13	9.8	10898	4	US-09-178-115-5	Sequence 5, Appli	c 960	12	9.1	21	1	US-09-517-467B-314	Sequence 314, App
c 888	13	9.8	10898	4	US-09-177-776-5	Sequence 5, Appli	c 961	12	9.1	21	1	US-08-151-574-36	Sequence 36, Appl
c 889	13	9.8	11298	1	US-07-869-933-31	Sequence 31, Appl	c 962	12	9.1	21	1	US-08-146-422-3	Sequence 3, Appli
c 890	13	9.8	11298	1	US-08-201-879A-2	Sequence 2, Appli	c 963	12	9.1	21	1	US-08-146-424-3	Sequence 3, Appli
c 891	13	9.8	11298	4	US-09-103-663-31	Sequence 31, Appl	c 964	12	9.1	21	1	US-08-626-554-19	Sequence 19, Appl
c 892	13	9.8	11464	4	US-08-591-840A-2	Sequence 2, Appli	c 965	12	9.1	21	1	US-08-693-709-17	Sequence 17, Appl
c 893	13	9.8	11873	2	US-08-970-269A-32	Sequence 32, Appl	c 966	12	9.1	21	1	US-08-638-448-3	Sequence 3, Appli
c 894	13	9.8	11873	4	US-09-407-562-32	Sequence 32, Appl	c 967	12	9.1	21	2	US-08-419-448-36	Sequence 36, Appl
c 895	13	9.8	11878	2	US-08-970-269A-31	Sequence 31, Appl	c 968	12	9.1	21	2	US-09-233-510-36	Sequence 36, Appl
c 896	13	9.8	11878	4	US-09-407-562-31	Sequence 31, Appl	c 969	12	9.1	22	2	US-08-505-617-10	Sequence 10, Appl
c 897	13	9.8	11883	2	US-08-970-269A-28	Sequence 28, Appl	c 970	12	9.1	22	3	US-09-018-170-10	Sequence 10, Appl
c 898	13	9.8	11883	4	US-09-407-562-28	Sequence 28, Appl	c 971	12	9.1	22	4	US-07-791-931-12	Sequence 12, Appl
c 899	13	9.8	12804	4	US-09-453-702B-149	Sequence 149, App	c 972	12	9.1	22	4	US-09-177-650-26	Sequence 26, Appl
c 900	13	9.8	13146	2	US-08-724-354D-3	Sequence 3, Appli	c 973	12	9.1	24	1	US-08-131-365B-46	Sequence 46, Appl
c 901	13	9.8	13146	3	US-09-270-984A-3	Sequence 3, Appli	c 974	12	9.1	24	1	US-08-302-537A-11	Sequence 11, Appl
c 902	13	9.8	13158	2	US-08-687-080-105	Sequence 105, App	c 975	12	9.1	24	1	US-08-579-667-10	Sequence 10, Appl
c 903	13	9.8	13965	4	US-09-453-702B-48	Sequence 48, Appl	c 976	12	9.1	24	1	US-08-566-037A-12	Sequence 12, Appl

c 977 12 9.1 24 2 US-08-568-123-46
c 978 12 9.1 24 2 US-08-987-466-10
c 979 12 9.1 24 2 US-08-260-546-4
c 980 12 9.1 24 2 US-08-974-565C-15
c 981 12 9.1 24 4 US-09-240-359-10
c 982 12 9.1 24 4 US-09-138-736-6
c 983 12 9.1 24 4 US-08-702-665A-14
c 984 12 9.1 24 4 US-09-436-068A-4
c 985 12 9.1 25 1 US-07-718-274A-38
c 986 12 9.1 25 1 US-08-149-106-38
c 987 12 9.1 25 2 US-08-298-021-38
c 988 12 9.1 25 2 US-08-596-319-18
c 989 12 9.1 27 3 US-08-997-803-9
c 990 12 9.1 27 3 US-09-005-165-9
c 991 12 9.1 27 4 US-08-584-040-5189
c 992 12 9.1 27 4 US-08-584-040-6419
c 993 12 9.1 28 1 US-08-336-526A-3
c 994 12 9.1 28 1 US-08-478-940-3
c 995 12 9.1 28 2 US-08-882-435-2
c 996 12 9.1 28 4 US-09-386-642-48
c 997 12 9.1 29 1 US-08-365-189-4
c 998 12 9.1 30 1 US-08-336-526A-4
c 999 12 9.1 30 1 US-08-478-940-4
c1000 12 9.1 30 1 US-08-450-945-34

ALIGNMENTS

RESULT 1
US-09-397-787-308/c
; Sequence 308, Application US/09397787
; Patent No. 6468758

; GENERAL INFORMATION:
; APPLICANT: Benson, Darin R.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: King, Gordon E.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR OVARIAN
; FILE REFERENCE: 210121.466C2
; CURRENT APPLICATION NUMBER: US/09/397,787
; CURRENT FILING DATE: 1999-09-16
; NUMBER OF SEQ ID NOS: 334
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 308
; LENGTH: 439
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-397-787-308

Query Match 12.9%; Score 17; DB 4; Length 439;
Best Local Similarity 100.0%; Pred. No. 7.1;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 50 ACCTGAAGGCCACACAG 66
Db 401 AGCTGAAGGCCACACAG 385
|||||

RESULT 2
US-09-484-970B-82/c
; Sequence 82, Application US/09484970B
; Patent No. 6426186
; GENERAL INFORMATION:
; APPLICANT: Jones, Karen A.
; APPLICANT: Volkmut, Wayne
; APPLICANT: Walker, Michael G.
; TITLE OF INVENTION: BONE REMODELING GENES
; FILE REFERENCE: PB-0014 US
; CURRENT APPLICATION NUMBER: US/09/484,970B
; CURRENT FILING DATE: 2000-01-18
; NUMBER OF SEQ ID NOS: 172

Sequence 46, Appl
Sequence 10, Appl
Sequence 4, Appl
Sequence 15, Appl
Sequence 10, Appl
Sequence 6, Appl
Sequence 14, Appl
Sequence 38, Appl
Sequence 38, Appl
Sequence 18, Appl
Sequence 9, Appl
Sequence 5189, Ap
Sequence 6419, Ap
Sequence 3, Appl
Sequence 2, Appl
Sequence 48, Appl
Sequence 4, Appl
Sequence 4, Appl
Sequence 4, Appl
Sequence 34, Appl

; SOFTWARE: PERL Program
; SEQ ID NO 82
; LENGTH: 776
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. 6426186 199069.2CB1
US-09-484-970B-82

Query Match 12.9%; Score 17; DB 4; Length 776;
Best Local Similarity 100.0%; Pred. No. 7.1;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 50 ACCTGAAGGCCACACAG 66
Db 691 AGCTGAAGGCCACACAG 675
|||||

RESULT 3
US-08-445-520B-8
; Sequence 8, Application US/08445520B
; Patent No. 5866323
; GENERAL INFORMATION:
; APPLICANT: Markowitz, Sanford D.
; APPLICANT: Brattain, Michael G.
; APPLICANT: Willison, James K.V.
; TITLE OF INVENTION: CANCER DIAGNOSIS, PROGNOSIS AND
; TITLE OF INVENTION: THERAPY BASED ON MUTATION OF RECEPTOR
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BAKER & BOTTS, L.L.P.
; STREET: 1299 Pennsylvania Avenue, N.W.
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20004-2400
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/445,520B
; FILING DATE: 22-MAY-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/417,867
; FILING DATE: 07-APR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Posorske, Laurence H
; REGISTRATION NUMBER: 34,698
; REFERENCE/DOCKET NUMBER: 062361-0101
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-639-7700
; TELEFAX: 202-639-7890
; TELEX:

INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 2090 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
FEATURE:
NAME/KEY: Coding Sequence
LOCATION: 336...2036
OTHER INFORMATION:
US-08-445-520B-8

Query Match 12.9%; Score 17; DB 2; Length 2090;
Best Local Similarity 100.0%; Pred. No. 7.2;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 41 ATAAGGCCCAAGCTGAAG 57
|||||
Db 1111 ATAAGGCCCAAGCTGAAG 1127

RESULT 4

US-08-451-946B-7
; Sequence 7, Application US/08451946B
; Patent No. 6001969
; GENERAL INFORMATION:
; APPLICANT: Lin, Herbert Y.
; APPLICANT: Wang, Xiao-Fan
; APPLICANT: Weinberg, Robert A.
; APPLICANT: Lodiish, Harvey F.
; TITLE OF INVENTION: TGF-Beta Type Receptor cDNAs Encoded
; TITLE OF INVENTION: Products and Uses Therefor
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
; STREET: Two Militia Drive
; CITY: Lexington
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02173

COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/451,946B
; FILING DATE: 26-MAY-1996
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/311,703
; FILING DATE: 23-SEP-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/786,063
; FILING DATE: 31-OCT-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Granahan, Patricia
; REGISTRATION NUMBER: 32,227
; REFERENCE/DOCKET NUMBER: WH191-09V
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-861-9540
; TELEFAX: 617-861-9540
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2090 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 336..2038
US-08-451-946B-7

Query Match 12.9%; Score 17; DB 3; Length 2090;
Best Local Similarity 100.0%; Pred. No. 7.2;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 41 ATAAGGCCCAAGCTGAAG 57
|||||
Db 1111 ATAAGGCCCAAGCTGAAG 1127

RESULT 5

US-08-446-938B-7
; Sequence 7, Application US/08446938B
; Patent No. 6008011
; GENERAL INFORMATION:
; APPLICANT: Lin, Herbert Y.

; APPLICANT: Wang, Xiao-Fan
; APPLICANT: Weinberg, Robert A.
; APPLICANT: Lodiish, Harvey F.
; TITLE OF INVENTION: TGF-Beta Type Receptor cDNAs Encoded
; TITLE OF INVENTION: Products and Uses Therefor
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
; STREET: Two Militia Drive
; CITY: Lexington
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02173

COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/446,938B
; FILING DATE: 23-MAY-1995
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/311,703
; FILING DATE: 23-SEP-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/786,063
; FILING DATE: 31-OCT-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Granahan, Patricia
; REGISTRATION NUMBER: 32,227
; REFERENCE/DOCKET NUMBER: WH191-09FZ
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-861-6240
; TELEFAX: 617-861-9540
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2090 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 336..2038
US-08-446-938B-7

Query Match 12.9%; Score 17; DB 3; Length 2090;
Best Local Similarity 100.0%; Pred. No. 7.2;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 41 ATAAGGCCCAAGCTGAAG 57
|||||
Db 1111 ATAAGGCCCAAGCTGAAG 1127

RESULT 6

US-08-311-703A-7
; Sequence 7, Application US/08311703A
; Patent No. 6010872
; GENERAL INFORMATION:
; APPLICANT: Lin, Herbert Y.

; APPLICANT: Wang, Xiao-Fan
; APPLICANT: Weinberg, Robert A.
; APPLICANT: Lodiish, Harvey F.
; TITLE OF INVENTION: TGF-Beta Type Receptor cDNAs Encoded
; TITLE OF INVENTION: Products and Uses Therefor
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
; STREET: Two Militia Drive
; CITY: Lexington
; STATE: Massachusetts

```
;; COUNTRY: U.S.A.
;; ZIP: 02173
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: Patentin Release #1.0, Version #1.30
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/311,703A
;; FILING DATE: 23-SEP-1994
;; CLASSIFICATION: 530
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 07/786,063
;; FILING DATE: 31-OCT-1991
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Hogle, Doreen
;; REGISTRATION NUMBER: 36,361
;; REFERENCE/DOCKET NUMBER: WHI91-09F
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 781-861-6240
;; TELEFAX: 781-861-9540
;; INFORMATION FOR SEQ ID NO: 7:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 2090 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: double
;; TOPOLOGY: linear
;; MOLECULE TYPE: DNA (genomic)
;; FEATURE:
;; NAME/KEY: CDS
;; LOCATION: 336..2038
US-08-311-703A-7
```

```
Query Match 12.9%; Score 17; DB 3; Length 2090;
Best Local Similarity 100.0%; Pred. No. 7.2;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 41 ATAAGGCCAAGCTGAAG 57
Db 1111 ATAAGGCCAAGCTGAAG 1127
```

```
RESULT 7
US-08-446-939B-7
; Sequence 7, Application US/08446939B
; Patent No. 6046157
; GENERAL INFORMATION:
; APPLICANT: Lin, Herbert Y.
; APPLICANT: Wang, Xiao-Fan
; APPLICANT: Weinberg, Robert A.
; APPLICANT: Lodish, Harvey F.
; TITLE OF INVENTION: TGF-Beta Type Receptor cDNAs Encoded
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
; STREET: Two Millia Drive
; CITY: Lexington
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02173
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/446,939B
; FILING DATE: 23-MAY-1995
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/311,703
; FILING DATE: 23-MAY-1995
```

```
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 07/786,063
;; FILING DATE: 23-MAY-1995
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Granahan, Patricia
;; REGISTRATION NUMBER: 32,227
;; REFERENCE/DOCKET NUMBER: WHI91-09FX
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 617-861-6240
;; TELEFAX: 617-861-9540
;; INFORMATION FOR SEQ ID NO: 7:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 2090 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: double
;; TOPOLOGY: linear
;; MOLECULE TYPE: DNA (genomic)
;; FEATURE:
;; NAME/KEY: CDS
;; LOCATION: 336..2038
US-08-446-939B-7
```

```
Query Match 12.9%; Score 17; DB 3; Length 2090;
Best Local Similarity 100.0%; Pred. No. 7.2;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 41 ATAAGGCCAAGCTGAAG 57
Db 1111 ATAAGGCCAAGCTGAAG 1127
```

```
RESULT 8
US-09-183-543-7
; Sequence 7, Application US/09183543A
; Patent No. 6086867
; GENERAL INFORMATION:
; APPLICANT: Lin, Herbert Y.
; APPLICANT: Wang, Xiao-Fan
; APPLICANT: Weinberg, Robert A.
; APPLICANT: Lodish, Harvey F.
; TITLE OF INVENTION: Modulation of TGF-Beta by TGF-Beta Type III Receptor Polypept
; FILE REFERENCE: WHI91-09FXA
; CURRENT APPLICATION NUMBER: US/09/183,543A
; CURRENT FILING DATE: 1998-10-30
; EARLIER APPLICATION NUMBER: US 08/446,939
; EARLIER FILING DATE: 1995-05-23
; EARLIER APPLICATION NUMBER: US 08/311,703
; EARLIER FILING DATE: 1994-09-23
; EARLIER APPLICATION NUMBER: US 07/786,063
; EARLIER FILING DATE: 1991-10-31
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 7
; LENGTH: 2090
; TYPE: DNA
; ORGANISM: genomic
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (336)...(2038)
US-09-183-543-7
```

```
Query Match 12.9%; Score 17; DB 3; Length 2090;
Best Local Similarity 100.0%; Pred. No. 7.2;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 41 ATAAGGCCAAGCTGAAG 57
Db 1111 ATAAGGCCAAGCTGAAG 1127
```

```
RESULT 9
US-08-446-936A-7
; Sequence 7, Application US/08446936A
```

Patent No. 6201108
; GENERAL INFORMATION:
; APPLICANT: Lin, Herbert Y.
; APPLICANT: Wang, Xiao-Fan
; APPLICANT: Weinberg, Robert A.
; APPLICANT: Lodish, Harvey F.
; TITLE OF INVENTION: TGF-Beta Type Receptor cDNAs Encoded
; TITLE OF INVENTION: Products and Uses Therefor
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
; STREET: Two Militia Drive
; CITY: Lexington
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02173
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/446,936A
; FILING DATE: 23-MAY-1995
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/311,703
; FILING DATE: 23-SEP-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/786,063
; FILING DATE: 31-OCT-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Granahan, Patricia
; REGISTRATION NUMBER: 32,227
; REFERENCE/DOCKET NUMBER: WH191-09FY
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-861-6240
; TELEFAX: 617-861-9540
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2090 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 336..2038
US-08-446-936A-7

Query Match 12.9%; Score 17; DB 4; Length 2090;
Best Local Similarity 100.0%; Pred. No. 7.2;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 41 ATAAGGCCCAAGCTGAAG 57
|||||
Db 1111 ATAAGGCCCAAGCTGAAG 1127

RESULT 10
US-09-239-864A-10
; Sequence 10, Application US/09239864A
; Patent No. 6291237
; GENERAL INFORMATION:
; APPLICANT: Markowitz, Sanford D
; APPLICANT: Brattain, Michael G
; APPLICANT: Willson, James K.V.
; TITLE OF INVENTION: CANCER DIAGNOSIS, PROGNOSIS AND THERAPY BASED ON
; TITLE OF INVENTION: MUTATION OF RECEPTOR
; FILE REFERENCE: 062361.0108
; CURRENT APPLICATION NUMBER: US/09/239,864A
; CURRENT FILING DATE: 1999-01-29
; PRIOR APPLICATION NUMBER: 08/417,867

; PRIOR FILING DATE: 1995-04-07
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO: 10
; LENGTH: 2090
; TYPE: DNA
; ORGANISM: human
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (336)..(2036)
US-09-239-864A-10

Query Match 12.9%; Score 17; DB 4; Length 2090;
Best Local Similarity 100.0%; Pred. No. 7.2;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 41 ATAAGGCCCAAGCTGAAG 57
|||||
Db 1111 ATAAGGCCCAAGCTGAAG 1127

RESULT 11
PCT-US92-09326-3
; Sequence 3, Application PC/TUS9209326
; GENERAL INFORMATION:
; APPLICANT: Whitehead Institute for Biomedical Research
; TITLE OF INVENTION: TGF-BETA TYPE RECEPTOR cDNAs ENCODED
; TITLE OF INVENTION: PRODUCTS AND USES THEREFOR
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
; STREET: Two Militia Drive
; CITY: Lexington
; STATE: MA
; COUNTRY: USA
; ZIP: 02173
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US92/09326
; FILING DATE: 19921030
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/786,063
; FILING DATE: 31-OCT-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Granahan, Patricia
; REGISTRATION NUMBER: 32,227
; REFERENCE/DOCKET NUMBER: WH191-09 PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 861-6240
; TELEFAX: (617) 861-9540
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2090 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
PCT-US92-09326-3

Query Match 12.9%; Score 17; DB 5; Length 2090;
Best Local Similarity 100.0%; Pred. No. 7.2;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 41 ATAAGGCCCAAGCTGAAG 57
|||||
Db 1111 ATAAGGCCCAAGCTGAAG 1127

RESULT 12
US-08-361-873A-1
; Sequence 1, Application US/08361873A
; Patent No. 5693607
; GENERAL INFORMATION:
; APPLICANT: SEGARINI, PATRICIA R.
; APPLICANT: DASCH, JAMES R.
; APPLICANT: OLSEN, DAVID R.
; APPLICANT: CARRILLO, PEDRO A.
; TITLE OF INVENTION: USES OF TGB-beta RECEPTOR FRAGMENT AS A
; TITLE OF INVENTION: THERAPEUTIC AGENT
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FORSTER
; STREET: 755 PAGE MILL ROAD
; CITY: PALO ALTO
; STATE: CA
; COUNTRY: USA
; ZIP: 94304-1018
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/361,873A
; FILING DATE: 22-DEC-1994
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: PARK, FREDDIE K.
; REGISTRATION NUMBER: 35,636
; REFERENCE/DOCKET NUMBER: 22095-20261.01
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 813-5600
; TELEFAX: (415) 494-0792
; TELEX: 706141
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2095 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 341..2041
US-08-361-873A-1
Query Match 12.9%; Score 17; DB 1; Length 2095;
Best Local Similarity 100.0%; Pred. No. 7.2;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 41 ATAAGGCCAAGCTGAAG 57
Db 1116 ATAAGGCCAAGCTGAAG 1132
RESULT 13
US-08-365-486A-12/c
; Sequence 12, Application US/08365486A
; Patent No. 5834306
; GENERAL INFORMATION:
; APPLICANT: Webster, Keith A.
; APPLICANT: Bishopric, Nanette H.
; TITLE OF INVENTION: Tissue Specific Hypoxia Regulated
; TITLE OF INVENTION: Therapeutic Constructs
; NUMBER OF SEQUENCES: 31
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dehlinger & Associates
; STREET: 350 Cambridge Avenue, Suite 250
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA

ZIP: 94306
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/365,486A
FILING DATE: 23-DEC-1994
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Sholtz, Charles K.
REGISTRATION NUMBER: 38,615
REFERENCE/DOCKET NUMBER: 8255-0018
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 324-0880
TELEFAX: (415) 324-0960
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 5057 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: CDNA to mRNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
INDIVIDUAL ISOLATE: rat bNOS CDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 349...4638
US-08-365-486A-12
Query Match 12.9%; Score 17; DB 2; Length 5057;
Best Local Similarity 100.0%; Pred. No. 7.3;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 32 CCAGCTTGGATAAGGCC 48
Db 530 CCAGCTTGGATAAGGCC 514
RESULT 14
US-08-880-342-12/c
; Sequence 12, Application US/08880342
; Patent No. 6218179
; GENERAL INFORMATION:
; APPLICANT: Webster, Keith A.
; APPLICANT: Bishopric, Nanette H.
; APPLICANT: Murphy, Brian
; APPLICANT: Laderoute, Keith R.
; APPLICANT: Green, Christopher J.
; TITLE OF INVENTION: Tissue Specific Hypoxia Regulated
; TITLE OF INVENTION: Therapeutic Constructs
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dehlinger & Associates
; STREET: 350 Cambridge Avenue, Suite 250
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/880,342
FILING DATE: 23-JUN-1997
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/IB95/00996

;; FILING DATE: 13-NOV-1995
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/365,486
;; FILING DATE: 23-DEC-1994
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Sholtz, Charles K.
;; REGISTRATION NUMBER: 38,615
;; REFERENCE/DOCKET NUMBER: 8255-0018.30
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (415) 324-0880
;; TELEFAX: (415) 324-0960
;; INFORMATION FOR SEQ ID NO: 12:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 5057 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: double
;; TOPOLOGY: linear
;; MOLECULE TYPE: cDNA to mRNA
;; HYPOTHETICAL: NO
;; ANTI-SENSE: NO
;; ORIGINAL SOURCE:
;; INDIVIDUAL ISOLATE: rat bNOS cDNA
;; FEATURE:
;; NAME/KEY: CDS
;; LOCATION: 349...4638
US-08-880-342-12

Query Match 12.9%; Score 17; DB 4; Length 5057;
Best Local Similarity 100.0%; Pred. No. 7.3;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 32 CCAGCTTGGATAAGGCC 48
|||||
DB 530 CCAGCTTGGATAAGGCC 514

RESULT 15
US-07-642-002-1/c
; Sequence 1, Application US/07642002
; Patent No. 5268465
; GENERAL INFORMATION:
; APPLICANT: Bredt, David S.
; APPLICANT: Hwang, Paul M.
; APPLICANT: Reed, Randall
; APPLICANT: Snyder, Solomon H.
; TITLE OF INVENTION: Purification and Molecular Cloning of Nitric
; TITLE OF INVENTION: Oxide Synthase
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Banner, Birch, McKie & Beckett
; STREET: One Thomas Circle, NW
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.24
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/642,002
; FILING DATE: 19910118
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Kagan, Sarah A.
; REGISTRATION NUMBER: 32,141
; REFERENCE/DOCKET NUMBER: 1107.033576
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 296-5500
; TELEFAX: (202) 296-7830
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:

;; LENGTH: 5108 base pairs
;; TYPE: NUCLEIC ACID
;; STRANDEDNESS: double
;; TOPOLOGY: linear
;; MOLECULE TYPE: cDNA
;; HYPOTHETICAL: N
;; ANTI-SENSE: N
;; ORIGINAL SOURCE:
;; ORGANISM: Rattus rattus
;; TISSUE TYPE: Brain
;; FEATURE:
;; NAME/KEY: CDS
;; LOCATION: 400...4686
;; OTHER INFORMATION:
US-07-642-002-1

Query Match 12.9%; Score 17; DB 1; Length 5108;
Best Local Similarity 100.0%; Pred. No. 7.3;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 32 CCAGCTTGGATAAGGCC 48
|||||
DB 581 CCAGCTTGGATAAGGCC 565

RESULT 16
US-08-700-637-9/c
; Sequence 9, Application US/08700637
; Patent No. 5854413
; GENERAL INFORMATION:
; APPLICANT: Hawkins, Phillip R.
; APPLICANT: Stuart, Susan G.
; APPLICANT: Murry, Lynn E.
; TITLE OF INVENTION: NOVEL SYNAPTOGVIRIN HOMOLOG FROM COLON
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: U.S.
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/700,637
; FILING DATE: Filed Herewith
; ATTORNEY/AGENT INFORMATION:
; NAME: Luther, Barbara J.
; REGISTRATION NUMBER: 33,954
; REFERENCE/DOCKET NUMBER: PF-0065 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-852-0195
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 220 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; IMMEDIATE SOURCE:
; LIBRARY: SYNORAT03
; CLONE: 695983
US-08-700-637-9

Query Match 12.1%; Score 16; DB 2; Length 220;
Best Local Similarity 100.0%; Pred. No. 23;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

: NAME: GRANADOS, Patricia D.
: REGISTRATION NUMBER: 33,683
: REFERENCE/DOCKET NUMBER: 26083/124
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (202)672-5300
: TELEFAX: (202)672-5399
: TELEX: 904136
: INFORMATION FOR SEQ ID NO: 9:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 3437 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: US-08-704-711A-9

Query Match 12.1%; Score 16
Best Local Similarity 100.0%; Pred. NC
Matches 16; Conservative 0; Mismatch 0

QY 52 CTGAGGCCACACAGA 67
|||||
Db 2701 CTGAGGCCACACAGA 2716

RESULT 19
US-09-521-220-9
: Sequence 9, Application US/09521220
: Patent No. 6399348
: GENERAL INFORMATION:
: APPLICANT: WILL, Horst
: HINZMANN, Bernd
: TITLE OF INVENTION: DNA SEQUENCES
: METALLOPROTEAS
: NUMBER OF SEQUENCES: 22
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Foley & Lardner
: STREET: 3000 K Street, N.W.,
: CITY: Washington
: STATE: D.C.
: COUNTRY: USA
: ZIP: 20007-5109
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-D
: SOFTWARE: PatentIn Release #1
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/521
: FILING DATE: 08-Mar-2000
: CLASSIFICATION: <Unknown>
: 21-OCT-1994
: 17-MAR-1994
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 08/704,71
: FILING DATE: <Unknown>
: APPLICATION NUMBER: DE 443883
: FILING DATE: 21-OCT-1994
: APPLICATION NUMBER: DE 440966
: FILING DATE: 17-MAR-1994
: ATTORNEY/AGENT INFORMATION:
: NAME: GRANADOS, Patricia D.
: REGISTRATION NUMBER: 33,683
: REFERENCE/DOCKET NUMBER: 2608
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (202)672-5300
: TELEFAX: (202)672-5399
: TELEX: 904136
: INFORMATION FOR SEQ ID NO: 9:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 3437 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear

```


; SEQUENCE DESCRIPTION: SEQ ID NO: 9;
US-09-521-220-9

Query Match 12.1%; Score 16; DB 4; Length 3437;
Best Local Similarity 100.0%; Pred. No. 24;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 52 CTGAAGGCCACAGAGA 67
|||||
Db 2701 CTGAAGGCCACAGAGA 2735

RESULT 20

US-08-704-711A-8
; Sequence 8, Application US/08704711A
; Patent No. 6114159

; GENERAL INFORMATION:
; APPLICANT: WILL, Horst
; TITLE OF INVENTION: DNA SEQUENCES FOR MATRIX
; TITLE OF INVENTION: METALLOPROTEASES, THEIR PRODUCTION AND USE
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20007-5109

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/704,711A
; FILING DATE: 20-NOV-1996
; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/DE95/00357
; FILING DATE: 17-MAR-1995
; PRIOR APPLICATION DATA: DE 4438838.1
; FILING DATE: 21-OCT-1994
; APPLICATION DATA:
; NAME: GRANADOS, Patricia D.
; REGISTRATION NUMBER: 33,683
; REFERENCE/DOCKET NUMBER: 26083/124
; TELEPHONE: (202)672-5300
; TELEFAX: (202)672-5399
; TELEX: 904136

; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3456 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear

; SEQUENCE DESCRIPTION: SEQ ID NO: 8:
US-08-704-711A-8

Query Match 12.1%; Score 16; DB 3; Length 3456;
Best Local Similarity 100.0%; Pred. No. 24;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 52 CTGAAGGCCACAGAGA 67
|||||
Db 2720 CTGAAGGCCACAGAGA 2735

RESULT 21

US-07-582-945-1/c
; Sequence 1, Application US/07582945
; Patent No. 5369019

; GENERAL INFORMATION:
; APPLICANT: TARKER FOGED, Niels
; APPLICANT: PETERSON, Svend
; TITLE OF INVENTION: A PASTURELLA VACCINE
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria

QY 52 CTGAAGGCCACAGAGA 67
|||||
Db 2720 CTGAAGGCCACAGAGA 2735

US-09-521-220-8
; Sequence 8, Application US/09521220
; Patent No. 6399348

; GENERAL INFORMATION:
; APPLICANT: WILL, Horst
; TITLE OF INVENTION: DNA SEQUENCES FOR MATRIX
; TITLE OF INVENTION: METALLOPROTEASES, THEIR PRODUCTION AND USE
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20007-5109

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/521,220
; FILING DATE: 08-Mar-2000
; CLASSIFICATION: <Unknown>
; 21-OCT-1994
; 17-MAR-1994

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/704,711
; FILING DATE: <Unknown>
; APPLICATION NUMBER: DE 4438838.1
; FILING DATE: 21-OCT-1994
; APPLICATION NUMBER: DE 4409663.1
; FILING DATE: 17-MAR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: GRANADOS, Patricia D.
; REGISTRATION NUMBER: 33,683
; REFERENCE/DOCKET NUMBER: 26083/124
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)672-5300
; TELEFAX: (202)672-5399
; TELEX: 904136

; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3456 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear

; SEQUENCE DESCRIPTION: SEQ ID NO: 8:
US-09-521-220-8

Query Match 12.1%; Score 16; DB 4; Length 3456;
Best Local Similarity 100.0%; Pred. No. 24;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 52 CTGAAGGCCACAGAGA 67
|||||
Db 2720 CTGAAGGCCACAGAGA 2735

RESULT 22

US-07-582-945-1/c
; Sequence 1, Application US/07582945
; Patent No. 5369019

; GENERAL INFORMATION:
; APPLICANT: TARKER FOGED, Niels
; APPLICANT: PETERSON, Svend
; TITLE OF INVENTION: A PASTURELLA VACCINE
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria

QY 52 CTGAAGGCCACAGAGA 67
|||||
Db 2720 CTGAAGGCCACAGAGA 2735

STATE: VA
COUNTRY: USA
ZIP: 22313-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/582.945
FILING DATE: 19901012
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 30307/112 PLVI
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)836-9300
TELEFAX: (703)683-4109
TELEX: 899149
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 4380 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
ORIGINAL SOURCE: Pasteurella multocida
FEATURE:
NAME/KEY: CDS
LOCATION: 219..4073
US-07-582-945-1

Query Match 12.1%; Score 16; DB 1; Length 4380;
Best Local Similarity 100.0%; Pred. No. 24;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 CTGGAAGAAATTGCCA 34
Db 3459 CTGGAAGAAATTGCCA 3444

RESULT 23
US-08-453-141-1/c
Sequence 1, Application US/08453141
Patent No. 5885589
GENERAL INFORMATION:
APPLICANT: FOGED, Niels T.
TITLE OF INVENTION: PASTEURILLA VACCINE
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 3000 K Street, N.W., Suite 500
CITY: Washington, D.C.
COUNTRY: USA
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/453.141
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/293.314
FILING DATE: 22-AUG-1994
APPLICATION NUMBER: US 07/582.945
FILING DATE: 12-OCT-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/582.945
FILING DATE: 12-OCT-1990
PRIOR APPLICATION DATA:

APPLICATION NUMBER: PCT/DR89/00084
FILING DATE: 11-APR-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: DK 1995/88
FILING DATE: 04-DEC-1988
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 40389/102/AKZO
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)672-5300
TELEFAX: (202)672-5399
TELEX: 904136
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 4380 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
FEATURE:
NAME/KEY: CDS
LOCATION: 219..4076
US-08-453-141-1

Query Match 12.1%; Score 16; DB 2; Length 4380;
Best Local Similarity 100.0%; Pred. No. 24;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 CTGGAAGAAATTGCCA 34
Db 3459 CTGGAAGAAATTGCCA 3444

RESULT 24
US-08-293-314-1/c
Sequence 1, Application US/08293314
Patent No. 6110470
GENERAL INFORMATION:
APPLICANT: FOGED, Niels T.
TITLE OF INVENTION: PASTEURILLA VACCINE
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 3000 K Street, N.W., Suite 500
CITY: Washington, D.C.
COUNTRY: USA
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/293.314
FILING DATE: 22-AUG-1994
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/582.945
FILING DATE: 12-OCT-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/DR89/00084
FILING DATE: 11-APR-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: DK 1995/88
FILING DATE: 04-DEC-1988
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 40389/102/AKZO
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)672-5300
TELEFAX: (202)672-5399

; TELEX: 904136
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4380 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 219..4076
US-08-293-314-1

Query Match 12.1%; Score 16; DB 3; Length 4380;
Best Local Similarity 100.0%; Pred. No. 24;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 CTGGAAGAAATTGCCA 34
|||||
DB 3459 CTGGAAGAAATTGCCA 3444

RESULT 25
US-08-253-155A-9
; Sequence 9, Application US/08253155A
; Patent No. 5691147
; GENERAL INFORMATION:
; APPLICANT: Gyuris, Jeno
; APPLICANT: Draetta, Giulio
; TITLE OF INVENTION: CDK4 Binding Proteins
; NUMBER OF SEQUENCES: 95
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 State Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII(text)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/253,155A
; FILING DATE: 02-JUN-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Vincent, Matthew P.
; REGISTRATION NUMBER: 36,709
; REFERENCE/DOCKET NUMBER: MII-028
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 227-5941
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8201 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-08-253-155A-9

Query Match 12.1%; Score 16; DB 1; Length 8201;
Best Local Similarity 100.0%; Pred. No. 24;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 82 CTGATGACCAAGAGA 97
|||||
DB 5803 CTGATGACCAAGAGA 5818

RESULT 26
US-09-453-702B-252/c

; Sequence 252, Application US/09453702B
; Patent No. 6365723
; GENERAL INFORMATION:
; APPLICANT: Blattner, Frederick R.
; Burland, Valerie
; Perna, Nicole T.
; Plunkett, Guy
; Welch, Rod
; TITLE OF INVENTION: No. 6365723el Sequences of E. coli O157
; NUMBER OF SEQUENCES: 265
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Quarles & Brady
; STREET: 1 South Pinckney Street
; CITY: Madison
; STATE: WI
; COUNTRY: US
; ZIP: 53701-2113
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch. 1.44Mb storage
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 8.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/453,702B
; FILING DATE: 03-Dec-1999
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/110,955
; FILING DATE: 04-DEC-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Seay, Nicholas J.
; REGISTRATION NUMBER: 27386
; REFERENCE/DOCKET NUMBER: 960296.95017
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (608) 251-5000
; TELEFAX: (608) 251-9166
; INFORMATION FOR SEQ ID NO: 252:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 12848
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; SEQUENCE DESCRIPTION: SEQ ID NO: 252:
US-09-453-702B-252
Query Match 12.1%; Score 16; DB 4; Length 12848;
Best Local Similarity 100.0%; Pred. No. 24;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 98 CCACAGACGAGGAGAA 113
|||||
DB 3068 CCACAGACGAGGAGAA 3053
RESULT 27
US-09-245-041-3
; Sequence 3, Application US/09245041
; Patent No. 6274339
; GENERAL INFORMATION:
; APPLICANT: Moore, K.
; APPLICANT: Nagle, D.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE DIAGNOSIS AND TREATMENT
; OF BODY WEIGHT DISORDERS INCLUDING OBESITY
; TITLE OF INVENTION: OF BODY WEIGHT DISORDERS INCLUDING OBESITY
; FILE REFERENCE: 7853-136
; CURRENT APPLICATION NUMBER: US/09/245,041
; CURRENT FILING DATE: 1999-02-05
; EARLIER APPLICATION NUMBER: 60/093,630
; EARLIER FILING DATE: 1998-07-21
; EARLIER APPLICATION NUMBER: 60/104,978
; EARLIER FILING DATE: 1998-10-20
; NUMBER OF SEQ ID NOS: 131
; SOFTWARE: FastSeq for Windows Version 3.0

```
; SEQ ID NO 3
; LENGTH: 17056
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-245-041-3

Query Match      12.1%; Score 16; DB 4; Length 17056;
Best Local Similarity 100.0%; Pred. No. 24;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      101 CAGACGAGGAGGAGTG 116
      |||
Db      7892 CAGACGAGGAGGAGTG 7907

RESULT 28
US-09-453-702B-79
; Sequence 79, Application US/09453702B
; Patent No. 6365723
; GENERAL INFORMATION:
; APPLICANT: Blattner, Frederick R.
; Burland, Valerie
; Perna, Nicole T.
; Plunkett, Guy
; Welch, Rod
; TITLE OF INVENTION: No. 6365723el Sequences of E. coli O157
; NUMBER OF SEQUENCES: 265
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Quarles & Brady
; STREET: 1 South Pinckney Street
; CITY: Madison
; STATE: WI
; COUNTRY: US
; ZIP: 53701-2113
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch. 1.44Mb storage
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 8.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/453,702B
; FILING DATE: 03-Dec-1999
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/110,955
; FILING DATE: 04-DEC-1998
; ADDRESS/AGENT INFORMATION:
; NAME: Seay, Nicholas J.
; REGISTRATION NUMBER: 27386
; REFERENCE/DOCKET NUMBER: 960296.95017
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (608) 251-5000
; TELEFAX: (608) 251-9166
; INFORMATION FOR SEQ ID NO: 79:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 38155
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; SEQUENCE DESCRIPTION: SEQ ID NO: 79:

US-09-453-702B-79

Query Match      12.1%; Score 16; DB 4; Length 38155;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      98 CCACAGAGCAGGAGAA 113
      |||
Db      32706 CCACAGAGCAGGAGAA 32721

RESULT 29
US-09-453-702B-137/c
; Sequence 137, Application US/09453702B
; Patent No. 6365723
; GENERAL INFORMATION:
; APPLICANT: Blattner, Frederick R.
; Burland, Valerie
; Perna, Nicole T.
; Plunkett, Guy
; Welch, Rod
; TITLE OF INVENTION: No. 6365723el Sequences of E. coli O157
; NUMBER OF SEQUENCES: 265
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Quarles & Brady
; STREET: 1 South Pinckney Street
; CITY: Madison
; STATE: WI
; COUNTRY: US
; ZIP: 53701-2113
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch. 1.44Mb storage
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 8.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/453,702B
; FILING DATE: 03-Dec-1999
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/110,955
; FILING DATE: 04-DEC-1998
; ADDRESS/AGENT INFORMATION:
; NAME: Seay, Nicholas J.
; REGISTRATION NUMBER: 27386
; REFERENCE/DOCKET NUMBER: 960296.95017
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (608) 251-5000
; TELEFAX: (608) 251-9166
; INFORMATION FOR SEQ ID NO: 137:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 48908
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; SEQUENCE DESCRIPTION: SEQ ID NO: 137:

US-09-453-702B-137

Query Match      12.1%; Score 16; DB 4; Length 48908;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      98 CCACAGAGCAGGAGAA 113
      |||
Db      14259 CCACAGAGCAGGAGAA 14244

RESULT 30
US-08-153-799-9
; Sequence 9, Application US/08153799
; Patent No. 5766883
; GENERAL INFORMATION:
; APPLICANT: Ballance, David J
; APPLICANT: Goodey, Andrew R
; TITLE OF INVENTION: Polypeptides
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: R Hain Swope, BOC Health Care Inc
; STREET: 100 Mountain Avenue
; CITY: Murray Hill
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07974
; COMPUTER READABLE FORM:
```

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/153,799
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/847975
FILING DATE: 06-MAR-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 8909916.2
FILING DATE: 29-APR-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/GB90/00650
FILING DATE: 26-APR-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/775952
FILING DATE: 29-OCT-1991
ATTORNEY/AGENT INFORMATION:
NAME: Swope, R Hain
REGISTRATION NUMBER: 24864
REFERENCE/DOCKET NUMBER: 92H832
TELECOMMUNICATION INFORMATION:
TELEPHONE: (908) 665 2400
TELEFAX: (908) 771 6159
TELEX: 219484
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 75 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
FEATURE:
NAME/KEY: CDS
LOCATION: 3..38
OTHER INFORMATION: /product= "encodes amino acids"
OTHER INFORMATION: 183-194 of HSA"
FEATURE:
NAME/KEY: misc.feature
LOCATION: 39..53
OTHER INFORMATION: /function= "KEX2 spacer"
FEATURE:
NAME/KEY: CDS
LOCATION: 54..74
OTHER INFORMATION: /function= "encodes amino acids 1-7"
OTHER INFORMATION: of uPA"
FEATURE:
NAME/KEY: misc.feature
LOCATION: 1..75
OTHER INFORMATION: /function= "LINKER 9"
US-08-153-799-9

Query Match 11.4%; Score 15; DB 1; Length 75;
Best Local Similarity 100.0%; Pred. No. 73;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 30 TCCACGCTTGGATAA 44
Db 35 TCCACGCTTGGATAA 49

RESULT 31
US-09-040-984-83
; Sequence 83, Application US/09040984
; Patent No. 6210883
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Wang, Tongtong

TITLE OF INVENTION: COMPOUNDS AND METHODS FOR DIAGNOSIS
TITLE OF INVENTION: OF LUNG CANCER
NUMBER OF SEQUENCES: 86
CORRESPONDENCE ADDRESS:
ADDRESSEE: SEED AND BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: WA
COUNTRY: USA
ZIP: 98104
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/040,984
FILING DATE: 18-MAR-1998
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Maki, David J.
REGISTRATION NUMBER: 31,392
REFERENCE/DOCKET NUMBER: 210121.456
TELECOMMUNICATION INFORMATION:
TELEPHONE: 206-622-4900
TELEFAX: 206-282-6031
TELEX:
INFORMATION FOR SEQ ID NO: 83:
SEQUENCE CHARACTERISTICS:
LENGTH: 460 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-09-040-984-83
Query Match 11.4%; Score 15; DB 4; Length 460;
Best Local Similarity 100.0%; Pred. No. 75;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 116 GGAGTCAAAATTCCT 130
Db 180 GGAGTCAAAATTCCT 194

RESULT 32
US-09-123-912-83
; Sequence 83, Application US/09123912A
; Patent No. 6312695
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Wang, Tongtong
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY OF LUNG CANCER
; FILE REFERENCE: 210121.455C1
; CURRENT APPLICATION NUMBER: US/09/123,912A
; CURRENT FILING DATE: 1998-07-27
; PRIOR APPLICATION NUMBER: 09/040.802
; PRIOR FILING DATE: 1998-03-18
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 83
; LENGTH: 460
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (104)
; OTHER INFORMATION: Where n is a, c, g or t
; NAME/KEY: modified_base
; LOCATION: (118)
; OTHER INFORMATION: Where n is a, c, g or t
; NAME/KEY: modified_base
; LOCATION: (172)
; OTHER INFORMATION: Where n is a, c, g or t

```
; NAME/KEY: modified_base
; LOCATION: (401)
; OTHER INFORMATION: Where n is a, c, g or t
; NAME/KEY: modified_base
; LOCATION: (422)
; OTHER INFORMATION: Where n is a, c, g or t
; NAME/KEY: modified_base
; LOCATION: (423)
; OTHER INFORMATION: Where n is a, c, g or t
; NAME/KEY: modified_base
; LOCATION: (444)
; OTHER INFORMATION: Where n is a, c, g or t
; NAME/KEY: modified_base
; LOCATION: (449)
; OTHER INFORMATION: Where n is a, c, g or t
US-09-123-912-83
```

```
Query Match 11.4%; Score 15; DB 4; Length 460;
Best Local Similarity 100.0%; Pred. No. 75;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 116 GGAGTGAATTCCT 130
      |||||
Db 180 GGAGTGAATTCCT 194
```

RESULT 33

```
US-09-643-597-83
; Sequence 83, Application US/09643597
; Patent No. 6426072
; GENERAL INFORMATION:
```

```
; APPLICANT: Wang, Tongtong
; APPLICANT: Fan, Liqun
; APPLICANT: Kalos, Michael D.
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Hosken, Nancy
; APPLICANT: Fanger, Gary R.
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Henderson, Robert A.
; APPLICANT: McNeill, Patricia D.
```

```
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.455C11
; CURRENT APPLICATION NUMBER: US/09/643,597
; CURRENT FILING DATE: 2000-08-21
; NUMBER OF SEQ ID NOS: 369
; SOFTWARE: FastSeq for Windows Version 3.0
```

```
; SEQ ID NO 83
; LENGTH: 460
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(460)
; OTHER INFORMATION: n = A,T,C or G
US-09-643-597-83
```

```
Query Match 11.4%; Score 15; DB 4; Length 460;
Best Local Similarity 100.0%; Pred. No. 75;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 116 GGAGTGAATTCCT 130
      |||||
Db 180 GGAGTGAATTCCT 194
```

RESULT 34

```
US-08-671-525B-9/c
; Sequence 9, Application US/08671525B
; Patent No. 5703220
; GENERAL INFORMATION:
```

```
; APPLICANT: Yamada, Tadataka
; APPLICANT: Gantz, Ira
; TITLE OF INVENTION: Genes Encoding Melanocortin Receptors
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
```

```
; ADDRESSEE: Harness, Dickey & Pierce, P.L.C.
; STREET: P.O. Box 828
; CITY: Bloomfield Hills
; STATE: MI
; COUNTRY: US
```

ZIP: 48303

```
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
```

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/671,525B

FILING DATE: June 27, 1996

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Smith, DeAnn F.

REGISTRATION NUMBER: 36683

REFERENCE/DOCKET NUMBER: 2115-000853DVB

TELECOMMUNICATION INFORMATION:

TELEPHONE: (810)641-1600

TELEFAX: (810)641-0270

INFORMATION FOR SEQ ID NO: 9:

SEQUENCE CHARACTERISTICS:

LENGTH: 975 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

HYPOTHETICAL: NO

ANTI-SENSE: NO

ORIGINAL SOURCE:

ORGANISM: Mouse

FEATURE:

NAME/KEY: CDS

LOCATION: 1..975

US-08-671-525B-9

```
Query Match 11.4%; Score 15; DB 1; Length 975;
Best Local Similarity 100.0%; Pred. No. 76;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 59 CCACAGAGATGCAGA 73
      |||||
Db 379 CCACAGAGATGCAGA 365
```

RESULT 35

US-08-672-109B-9/c

Sequence 9, Application US/08672109B

Patent No. 5710265

GENERAL INFORMATION:

APPLICANT: Yamada, Tadataka

APPLICANT: Gantz, Ira

TITLE OF INVENTION: Genes Encoding Melanocortin Receptors

NUMBER OF SEQUENCES: 23

CORRESPONDENCE ADDRESS:

ADDRESSEE: Harness, Dickey & Pierce, P.L.C.

STREET: P.O. Box 828

CITY: Bloomfield Hills

STATE: MI

COUNTRY: US

ZIP: 48303

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/672,109B
;; FILING DATE: June 27, 1996
;; CLASSIFICATION: 536
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Smith, Deann F.
;; REGISTRATION NUMBER: 36683
;; REFERENCE/DOCKET NUMBER: 2115-000853DVC
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (810)641-1600
;; TELEFAX: (810)641-0270
;; INFORMATION FOR SEQ ID NO: 9:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 975 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: double
;; TOPOLOGY: linear
;; MOLECULE TYPE: DNA (genomic)
;; HYPOTHETICAL: NO
;; ANTI-SENSE: NO
;; ORIGINAL SOURCE:
;; ORGANISM: Mouse
;; FEATURE:
;; NAME/KEY: CDS
;; LOCATION: 1..975
US-08-672-109B-9

Query Match 11.4%; Score 15; DB 1; Length 975;
Best Local Similarity 100.0%; Pred. No. 76;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 59 CCACAGAGATGCAGA 73
|||||
DB 379 CCACAGAGATGCAGA 365

RESULT 36
US-08-842-045-9/c
;; Sequence 9, Application US/08842045
;; Patent No. 5817787
;; GENERAL INFORMATION:
;; APPLICANT: Yamada, Tadataka
;; APPLICANT: Gantz, Ira
;; TITLE OF INVENTION: Genes Encoding Melanocortin Receptors
;; NUMBER OF SEQUENCES: 23
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Harness, Dickey & Pierce, P.L.C.
;; STREET: P.O. Box 828
;; CITY: Bloomfield Hills
;; STATE: MI
;; COUNTRY: US
;; ZIP: 48303
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: Patent In Release #1.0, Version #1.25
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/842,045
;; FILING DATE:
;; CLASSIFICATION: 536
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Smith, Deann F.
;; REGISTRATION NUMBER: 36683
;; REFERENCE/DOCKET NUMBER: 2115-000853DVE
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (810)641-1600
;; TELEFAX: (810)641-0270
;; INFORMATION FOR SEQ ID NO: 9:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 975 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: double

;; TOPOLOGY: linear
;; MOLECULE TYPE: DNA (genomic)
;; HYPOTHETICAL: NO
;; ANTI-SENSE: NO
;; ORIGINAL SOURCE:
;; ORGANISM: Mouse
;; FEATURE:
;; NAME/KEY: CDS
;; LOCATION: 1..975
US-08-842-045-9

Query Match 11.4%; Score 15; DB 1; Length 975;
Best Local Similarity 100.0%; Pred. No. 76;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 59 CCACAGAGATGCAGA 73
|||||
DB 379 CCACAGAGATGCAGA 365

RESULT 37
US-08-842-238-9/c
;; Sequence 9, Application US/08842238
;; Patent No. 5869257
;; GENERAL INFORMATION:
;; APPLICANT: Yamada, Tadataka
;; APPLICANT: Gantz, Ira
;; TITLE OF INVENTION: Genes Encoding Melanocortin Receptors
;; NUMBER OF SEQUENCES: 23
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Harness, Dickey & Pierce, P.L.C.
;; STREET: P.O. Box 828
;; CITY: Bloomfield Hills
;; STATE: MI
;; COUNTRY: US
;; ZIP: 48303
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: Patent In Release #1.0, Version #1.25
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/842,238
;; FILING DATE:
;; CLASSIFICATION: 435
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Smith, Deann F.
;; REGISTRATION NUMBER: 36683
;; REFERENCE/DOCKET NUMBER: 2115-000853DVD
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (810)641-1600
;; TELEFAX: (810)641-0270
;; INFORMATION FOR SEQ ID NO: 9:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 975 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: double
;; TOPOLOGY: linear
;; MOLECULE TYPE: DNA (genomic)
;; HYPOTHETICAL: NO
;; ANTI-SENSE: NO
;; ORIGINAL SOURCE:
;; ORGANISM: Mouse
;; FEATURE:
;; NAME/KEY: CDS
;; LOCATION: 1..975
US-08-842-238-9

Query Match 11.4%; Score 15; DB 2; Length 975;
Best Local Similarity 100.0%; Pred. No. 76;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 59 CCACAGAGATGCAGA 73

TITLE OF INVENTION: Using Mammalian Melanocortin Receptor Agonists and Antagonists to Modulate Feeding Behavior in Animals

NUMBER OF SEQUENCES: 19

CORRESPONDENCE ADDRESS: McDonnell Boehnen Hulbert & Berghoff

STREET: 300 South Wacker Drive

CITY: Chicago

STATE: IL

COUNTRY: USA

ZIP: 60606

COMPUTER READABLE FORM: FLOPPY disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.25

CURRENT APPLICATION DATA: US/08/706,281A

FILING DATE: 04-SEP-1996

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION: NAME: NO. 6100048han, Kevin E

REGISTRATION NUMBER: 35,303

REFERENCE/DOCKET NUMBER: 96,886

TELEPHONE: 312-913-0001

TELEFAX: 312-913-0002

TELEX: INFORMATION FOR SEQ ID NO: 17:

SEQUENCE CHARACTERISTICS: LENGTH: 978 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

FEATURE: NAME/KEY: CDS

LOCATION: 1..975

US-08-706-281A-17

Query Match 11.4%; Score 15; DB 3; Length 978;

Best Local Similarity 100.0%; Pred. No. 76;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 59 CCACAGAGATGCAGA 73

Db 379 CCACAGAGATGCAGA 365

RESULT 40

US-09-097-231-17/C

Sequence 17, Application US/09097231

Patent No. 6278038

GENERAL INFORMATION: APPLICANT: Cone, Roger D

Chen, Wenbiao

Low, Malcolm J

TITLE OF INVENTION: Mammalian Melanocortin Receptor and Uses

NUMBER OF SEQUENCES: 22

CORRESPONDENCE ADDRESS: ADDRESSEE: McDonnell Boehnen Hulbert & Berghoff

STREET: 300 South Wacker Drive

CITY: Chicago

STATE: Illinois

COUNTRY: USA

ZIP: 60606

COMPUTER READABLE FORM: MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.25

CURRENT APPLICATION DATA: APPLICATION NUMBER: US/09/097,231

FILING DATE: 12-Jun-1998

Db 379 CCACAGATGCAGA 365

RESULT 38

US-08-629-335B-9/C

Sequence 9, Application US/08629335B

Patent No. 6117975

GENERAL INFORMATION: APPLICANT: Yamada, Tadataka

APPLICANT: Gantz, Ira

TITLE OF INVENTION: Genes Encoding Melanocortin Receptors

NUMBER OF SEQUENCES: 23

CORRESPONDENCE ADDRESS: ADDRESSEE: Harness, Dickey & Pierce, P.L.C.

STREET: P.O. Box 828

CITY: Bloomfield Hills

STATE: MI

COUNTRY: US

ZIP: 48303

COMPUTER READABLE FORM: FLOPPY disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.25

CURRENT APPLICATION DATA: APPLICATION NUMBER: US/08/629,335B

FILING DATE: July 23, 1996

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION: NAME: Smith, DeAnn F.

REGISTRATION NUMBER: 36683

REFERENCE/DOCKET NUMBER: 2115-000853DVA

TELEPHONE: (810)641-1600

TELEFAX: (810)641-0270

INFORMATION FOR SEQ ID NO: 9:

SEQUENCE CHARACTERISTICS: LENGTH: 975 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

HYPOTHETICAL: NO

ANTI-SENSE: NO

ORIGINAL SOURCE: ORGANISM: Mouse

FEATURE: NAME/KEY: CDS

LOCATION: 1..975

US-08-629-335B-9

Query Match 11.4%; Score 15; DB 3; Length 975;

Best Local Similarity 100.0%; Pred. No. 76;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 59 CCACAGATGCAGA 73

Db 379 CCACAGATGCAGA 365

RESULT 39

US-08-706-281A-17/C

Sequence 17, Application US/08706281A

Patent No. 6100048

GENERAL INFORMATION: APPLICANT: Cone, Roger D

APPLICANT: Fan, Wei

APPLICANT: Boston, Bruce A

APPLICANT: Kesterton, Robert A

APPLICANT: Lu, Dongsi

APPLICANT: Chen, Wenbiao

TITLE OF INVENTION: Methods and Reagents for Discovering and

CLASSIFICATION: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: No. 6278038nan, Kevin E
REGISTRATION NUMBER: 35,303
REFERENCE/DOCKET NUMBER: 96,886-C
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-913-0001
TELEFAX: 312-913-0002
TELEX: <Unknown>
INFORMATION FOR SEQ ID NO: 17:
SEQUENCE CHARACTERISTICS:
LENGTH: 978 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 1..975
SEQUENCE DESCRIPTION: SEQ ID NO: 17:
US-09-097-231-17

Query Match 11.4%; Score 15; DB 4; Length 978;
Best Local Similarity 100.0%; Pred. No. 76;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 59 CCACAGATGCAGA 73
Db 379 CCACAGATGCAGA 365
|||||

RESULT 41
US-09-152-060-12
; Sequence 12, Application US/09152060
; Patent No. 6448230
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 28 Human Secreted Proteins
; FILE REFERENCE: P2003P1.US
; CURRENT APPLICATION NUMBER: US/09/152,060
; CURRENT FILING DATE: 1998-09-11
; EARLIER APPLICATION NUMBER: PCT/US98/04858
; EARLIER FILING DATE: 1998-03-12
; EARLIER APPLICATION NUMBER: 60/040,762
; EARLIER FILING DATE: 1997-03-14
; EARLIER APPLICATION NUMBER: 60/040,710
; EARLIER FILING DATE: 1997-03-14
; EARLIER APPLICATION NUMBER: 60/050,934
; EARLIER FILING DATE: 1997-05-30
; EARLIER APPLICATION NUMBER: 60/048,100
; EARLIER FILING DATE: 1997-05-30
; EARLIER APPLICATION NUMBER: 60/048,357
; EARLIER FILING DATE: 1997-05-30
; EARLIER APPLICATION NUMBER: 60/048,189
; EARLIER FILING DATE: 1997-05-30
; EARLIER APPLICATION NUMBER: 60/057,765
; EARLIER FILING DATE: 1997-09-05
; EARLIER APPLICATION NUMBER: 60/048,970
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/068,368
; EARLIER FILING DATE: 1997-12-19
; NUMBER OF SEQ ID NOS: 118
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 12
; LENGTH: 1586
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-152-060-12

Query Match 11.4%; Score 15; DB 4; Length 1586;
Best Local Similarity 100.0%; Pred. No. 76;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 106 CAGGAGAAGTGGAGT 120
Db 11 CAGGAGAAGTGGAGT 25
|||||

RESULT 42
US-09-453-702B-105
; Sequence 105, Application US/09453702B
; Patent No. 6365723
; GENERAL INFORMATION:
; APPLICANT: Blattner, Frederick R.
; Burland, Valerie
; Perna, Nicole T.
; Plunkett, Guy
; Welch, Rod
; TITLE OF INVENTION: No. 6365723el Sequences of E. coli O157
; NUMBER OF SEQUENCES: 265
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Quarles & Brady
; STREET: 1 South Pinckney Street
; CITY: Madison
; STATE: WI
; COUNTRY: US
; ZIP: 53701-2113
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch. 1.44Mb storage
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 8.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/453,702B
; FILING DATE: 03-Dec-1999
; CLASSIFICATION: <Unknown>
; APPLICATION NUMBER: 60/110,955
; FILING DATE: 04-DEC-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Seay, Nicholas J.
; REGISTRATION NUMBER: 27386
; REFERENCE/DOCKET NUMBER: 960296.95017
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (608) 251-5000
; TELEFAX: (608) 251-9166
; INFORMATION FOR SEQ ID NO: 105:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1711
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; SEQUENCE DESCRIPTION: SEQ ID NO: 105:
US-09-453-702B-105

Query Match 11.4%; Score 15; DB 4; Length 1711;
Best Local Similarity 100.0%; Pred. No. 76;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 17 ACCTGGAAGAAATTG 31
Db 1489 ACCTGGAAGAAATTG 1503
|||||

RESULT 43
US-08-615-170-18
; Sequence 18, Application US/08615170
; Patent No. 5776776
; GENERAL INFORMATION:
; APPLICANT: ORDAHL, Charles P.
; APPLICANT: AZAKIE, Anthony
; APPLICANT: MAR, Janet H.
; APPLICANT: FARRANCE, Iain K.G.
; APPLICANT: HALL, Deborah E.
; APPLICANT: STEWART, Alexandre F.R.

APPLICANT: LARKIN, Sarah B.
TITLE OF INVENTION: DTEF-1 ISOFORMS AND USES THEREOF
NUMBER OF SEQUENCES: 32
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend Kourie and Crew
STREET: Steuart Street Tower, One Market Plaza
CITY: San Francisco
STATE: California
COUNTRY: US
ZIP: 94105-1493
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/615,170
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/191,493
FILING DATE: 04-FEB-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Heslin, James M.
REGISTRATION NUMBER: 29,541
REFERENCE/DOCKET NUMBER: 23070-053120
TELEPHONE: (415) 326-2400
TELEFAX: (415) 326-2422
INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 1897 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-08-615-170-18

Query Match 11.4%; Score 15; DB 1: Length 1897;
Best Local Similarity 100.0%; Pred. No. 76;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 47 CCAAGCTGAAGGCCA 61
|||||
Db 672 CCAAGCTGAAGGCCA 686

RESULT 44
US-09-434-613-2/c
Sequence 2, Application US/09434613
Patent No. 6337187
GENERAL INFORMATION:
APPLICANT: Kapeller-Libermann, Rosana
TITLE OF INVENTION: 18991, A No. 6337187el Human Lipase
FILE REFERENCE: 5800-63
CURRENT APPLICATION NUMBER: US/09/434,613
NUMBER OF SEQ ID NOS: 2
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 2
LENGTH: 1964
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (164)...(1174)
US-09-434-613-2

Query Match 11.4%; Score 15; DB 4; Length 1964;
Best Local Similarity 100.0%; Pred. No. 77;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 51 GCTGAAGGCCACAGA 65
|||||
Db 261 GCTGAAGGCCACAGA 247

RESULT 45
US-09-149-476-132/c
Sequence 132, Application US/09149476
Patent No. 6420526
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: 186 Human Secreted proteins
FILE REFERENCE: P2002P1
CURRENT APPLICATION NUMBER: US/09/149,476
CURRENT FILING DATE: 1998-09-08
EARLIER APPLICATION NUMBER: PCT/US98/04493
EARLIER FILING DATE: 1998-03-06
EARLIER APPLICATION NUMBER: 60/040,162
EARLIER FILING DATE: 1997-03-07
EARLIER APPLICATION NUMBER: 60/040,333
EARLIER FILING DATE: 1997-03-07
EARLIER APPLICATION NUMBER: 60/038,621
EARLIER FILING DATE: 1997-03-07
EARLIER APPLICATION NUMBER: 60/040,626
EARLIER FILING DATE: 1997-03-07
EARLIER APPLICATION NUMBER: 60/040,334
EARLIER FILING DATE: 1997-03-07
EARLIER APPLICATION NUMBER: 60/040,336
EARLIER FILING DATE: 1997-03-07
EARLIER APPLICATION NUMBER: 60/040,163
EARLIER FILING DATE: 1997-03-07
EARLIER APPLICATION NUMBER: 60/047,600
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,615
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,597
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,502
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,633
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,583
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,617
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,618
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,503
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,592
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,581
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,584
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,500
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,587
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,492
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,598
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,613
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,582
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,596
EARLIER FILING DATE: 1997-05-23

EARLIER APPLICATION NUMBER: 60/047,612
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,632
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,601
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/043,580
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,568
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,314
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,569
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,311
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,671
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,674
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,669
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,312
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,313
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,672
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,315
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/048,974
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/056,886
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,877
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,889
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,893
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,630
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,878
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,662
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,872
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,882
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,637
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,903
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,888
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,879
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,880
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,894
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,911
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,636
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,874
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,910
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,864
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,631

EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,845
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,892
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/057,761
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/047,595
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,599
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,588
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,585
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,586
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,590
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,594
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,589
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,593
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,614
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/043,578
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,576
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/047,501
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/043,670
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/056,632
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,664
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,876
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,881
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,909
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,875
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,862
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,887
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,908
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/048,964
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/057,650
EARLIER FILING DATE: 1997-09-05
EARLIER APPLICATION NUMBER: 60/056,884
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/057,669
EARLIER FILING DATE: 1997-09-05
EARLIER APPLICATION NUMBER: 60/049,610
EARLIER FILING DATE: 1997-06-13
EARLIER APPLICATION NUMBER: 60/061,060
EARLIER FILING DATE: 1997-10-02

Query Match 11.4%; Score 15; DB 4; Length 2012;
Best Local Similarity 100.0%; Pred. No. 77;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 118 AGTGAATTTCTGA 132
DB 1688 AGTGAATTTCTGA 1674
|||||

RESULT 47
US-08-494-714-1
; Sequence 1, Application US/08494714
; Patent No. 5587290
; GENERAL INFORMATION:
; APPLICANT: Kilonsky, Daniel.
; APPLICANT: Holzer, Helmut
; APPLICANT: Destruelle, Monica
; TITLE OF INVENTION: STRESS TOLERANT YEAST MUTANTS
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESS: FLEHR, HOHBACH, TEST, ALBRITTON & HE
; STREET: 4 Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-4187
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/494,714
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Osman Ph.D., Richard Aron
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: A-61036/DVB/RAO

Query Match 11.4%; Score 15; DB 5; Length 3876;
Best Local Similarity 100.0%; Pred. No. 77;
Matches 15; Conservative 0; Mismatches 0; Indels

QY 115 TGGAGTGAATTC 129
|||||
Db 2229 TGGAGTGAATTC 2243

RESULT 49

US-09-295-593-37
; Sequence 37, Application US/09295593
; Patent No. 6417169
; GENERAL INFORMATION:
; APPLICANT: WRIGHT, Jim A.
; APPLICANT: YOUNG, Alping H.
; APPLICANT: LEE, Yoon S.
; TITLE OF INVENTION: INSULIN-LIKE GROWTH FACTOR II ANTISENSE OLIGONUCLEOTIDE
; TITLE OF INVENTION: SEQUENCES AND METHODS OF USING SAME TO MODULATE CELL
; TITLE OF INVENTION: GROWTH
; FILE REFERENCE: 032396-046
; CURRENT APPLICATION NUMBER: US/09/295,593
; CURRENT FILING DATE: 1999-04-22
; EARLIER APPLICATION NUMBER: US 60/082,791
; EARLIER FILING DATE: 1998-04-23
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 37
; LENGTH: 4350
; TYPE: DNA
; ORGANISM: Human
US-09-295-593-37

Query Match 11.4%; Score 15; DB 4; Length 4350;
Best Local Similarity 100.0%; Pred. No. 77;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 98 CCACAGACGAGGAGA 112
|||||
Db 3584 CCACAGACGAGGAGA 3598

RESULT 50

US-09-066-047-1/c
; Sequence 1, Application US/09066047A
; Patent No. 6306394

GENERAL INFORMATION:

APPLICANT: MURPHY, Cheryl

STOREY, James

BELTZ, Gerald A.

COUGHLIN, Richard T.

TITLE OF INVENTION: NUCLEIC ACIDS, PROTEINS, AND METHODS OF
USE OF GRANULOCYTIC ERHLLCHIA

NUMBER OF SEQUENCES: 41

CORRESPONDENCE ADDRESS:

ADDRESSER: HALE AND DORR LLP

STREET: 60 State Street

CITY: Boston

STATE: Massachusetts

COUNTRY: United States

ZIP: 02109

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/066,047A

FILING DATE: 24-Apr-1998

CLASSIFICATION: <unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 60/044,869

FILING DATE: 25-APR-1997

ATTORNEY/AGENT INFORMATION:

NAME: Superko, Colleen

REGISTRATION NUMBER: 39,850

REFERENCE/DOCKET NUMBER: 106.941.156

; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 526-6000
; TELEFAX: (617) 526-5000
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4833 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-066-047-1

Query Match 11.4%; Score 15; DB 4; Length 4833;
Best Local Similarity 100.0%; Pred. No. 78;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 71 AGAAGACACTCTGA 85
|||||
Db 4006 AGAAGACACTCTGA 3992

Search completed: June 2, 2003, 16:39:48
Job time : 109 secs

THIS PAGE BLANK (3PT0)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 2, 2003, 15:56:19 ; Search time 106 Seconds
(without alignments)
1679.096 Million cell updates/sec

Title: US-09-915-178-1

Perfect score: 132

Sequence: 1 atggcacacaaactagacct.....agtggagtgaattctcta 132

Scoring table: OLIGO_NUC

Gapop 60.0 , Gapext 60.0

Searched: 845702 seqs, 674182571 residues

Word size : 10

Total number of hits satisfying chosen parameters: 199495

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database : Published Applications,NA:*
1: /cgn2.6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2.6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2.6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2.6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2.6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2.6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2.6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2.6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2.6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
10: /cgn2.6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
11: /cgn2.6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
12: /cgn2.6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
13: /cgn2.6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
14: /cgn2.6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	19	14.4	39776	9	US-10-160-293-3
2	18	13.6	243	10	US-09-919-580-517
3	18	13.6	353	10	US-09-919-580-224
4	18	13.6	395	10	US-09-919-580-207
5	18	13.6	404	10	US-09-960-352-10885
6	18	13.6	405	10	US-09-960-352-6906
7	18	13.6	408	10	US-09-960-352-12808
8	18	13.6	426	10	US-09-998-598-71
9	18	13.6	428	10	US-09-919-580-503
10	18	13.6	429	10	US-09-919-580-633
11	18	13.6	434	10	US-09-919-580-627
12	18	13.6	438	10	US-09-919-580-458
13	18	13.6	438	10	US-09-919-580-583
14	18	13.6	439	10	US-09-919-580-910
15	18	13.6	445	10	US-09-919-580-659
16	18	13.6	445	10	US-09-919-580-727
17	18	13.6	446	10	US-09-919-580-891
18	18	13.6	446	10	US-09-880-107-3158
19	18	13.6	449	9	US-09-918-995-15768

20	18	13.6	449	10	US-09-919-580-122	Sequence 122, App
21	18	13.6	451	10	US-09-919-580-190	Sequence 190, App
22	18	13.6	451	10	US-09-919-580-553	Sequence 553, App
23	18	13.6	452	10	US-09-919-580-348	Sequence 348, App
24	18	13.6	453	10	US-09-919-580-274	Sequence 274, App
25	18	13.6	453	10	US-09-880-107-3019	Sequence 3019, App
26	18	13.6	455	10	US-09-919-580-273	Sequence 273, App
27	18	13.6	456	10	US-09-919-580-219	Sequence 219, App
28	18	13.6	461	10	US-09-960-352-2605	Sequence 2605, App
29	18	13.6	465	10	US-09-919-580-56	Sequence 56, Appl
30	18	13.6	468	9	US-09-918-995-21161	Sequence 21161, A
31	18	13.6	469	9	US-09-918-995-24062	Sequence 24062, A
32	18	13.6	473	9	US-09-918-995-15273	Sequence 15273, A
33	18	13.6	474	9	US-09-918-995-22477	Sequence 22477, A
34	18	13.6	483	9	US-09-918-995-19906	Sequence 19906, A
35	18	13.6	494	9	US-09-918-995-24178	Sequence 24178, A
36	18	13.6	517	10	US-09-919-580-874	Sequence 874, App
37	18	13.6	559	10	US-09-919-580-159	Sequence 159, App
38	18	13.6	693	10	US-09-919-580-798	Sequence 798, App
39	18	13.6	9953	9	US-10-092-154-1336	Sequence 1336, Ap
40	18	13.6	9953	10	US-09-764-847-1336	Sequence 1336, Ap
41	18	13.6	11189	9	US-10-092-154-1335	Sequence 1335, Ap
42	18	13.6	11189	10	US-09-764-847-1335	Sequence 1335, Ap
43	17	12.9	425	9	US-09-918-995-6168	Sequence 6168, Ap
44	17	12.9	428	10	US-09-960-352-5941	Sequence 5941, Ap
45	17	12.9	433	9	US-09-918-995-4250	Sequence 4250, Ap
46	17	12.9	439	10	US-09-876-889-308	Sequence 308, App
47	17	12.9	461	9	US-09-918-995-22450	Sequence 22450, A
48	17	12.9	465	9	US-09-918-995-23067	Sequence 23067, A
49	17	12.9	814	9	US-10-102-806-146	Sequence 146, App
50	17	12.9	1077	9	US-09-738-626-1039	Sequence 1039, Ap
51	17	12.9	2090	10	US-09-878-905-10	Sequence 10, Appl
52	17	12.9	5759	10	US-09-954-456-2001	Sequence 2001, Ap
53	17	12.9	5821	9	US-10-071-766-52	Sequence 52, Appl
54	17	12.9	13026	9	US-09-764-891-10217	Sequence 10217, A
55	17	12.9	3309400	9	US-09-738-626-1	Sequence 1, Appl
56	16	12.1	134	9	US-09-925-299-674	Sequence 674, App
57	16	12.1	134	10	US-09-925-299-674	Sequence 674, App
58	16	12.1	281	9	US-10-060-036-1389	Sequence 1389, Ap
59	16	12.1	336	10	US-09-983-965-1398	Sequence 1398, Ap
60	16	12.1	376	10	US-09-878-574-408	Sequence 408, App
61	16	12.1	434	9	US-10-092-154-1925	Sequence 1925, Ap
62	16	12.1	444	10	US-09-764-847-1925	Sequence 1925, Ap
63	16	12.1	444	10	US-09-962-436-44	Sequence 44, Appl
64	16	12.1	444	10	US-09-880-107-131	Sequence 131, App
65	16	12.1	460	9	US-09-925-299-674	Sequence 674, App
66	16	12.1	460	10	US-09-764-847-466	Sequence 466, App
67	16	12.1	475	10	US-09-864-761-12	Sequence 12, Appl
68	16	12.1	475	10	US-09-864-761-101	Sequence 101, Appl
69	16	12.1	478	9	US-09-918-995-35268	Sequence 35268, A
70	16	12.1	479	10	US-09-867-550-1367	Sequence 1367, Ap
71	16	12.1	532	9	US-09-918-995-26427	Sequence 26427, A
72	16	12.1	665	9	US-10-103-313-122	Sequence 122, App
73	16	12.1	665	9	US-10-115-928-22	Sequence 22, Appl
74	16	12.1	870	10	US-09-815-242-6650	Sequence 6650, Ap
75	16	12.1	977	9	US-10-016-157A-105	Sequence 105, App
76	16	12.1	10385	10	US-09-864-761-18950	Sequence 18950, A
77	16	12.1	1385	9	US-10-097-340-86	Sequence 86, Appl
78	16	12.1	1453	10	US-09-349-015-14	Sequence 14, Appl
79	16	12.1	1492	10	US-09-864-761-2209	Sequence 2209, Ap
80	16	12.1	1584	9	US-10-103-313-636	Sequence 636, App
81	16	12.1	1584	9	US-10-115-928-68	Sequence 68, Appl
82	16	12.1	2199	10	US-09-381-624A-5	Sequence 5, Appl
83	16	12.1	2355	10	US-09-745-763-99	Sequence 99, Appl
84	16	12.1	2693	9	US-10-103-313-635	Sequence 635, App
85	16	12.1	2693	9	US-10-103-313-637	Sequence 637, App
86	16	12.1	2693	9	US-10-103-313-638	Sequence 638, App
87	16	12.1	2693	9	US-10-115-928-67	Sequence 67, Appl
88	16	12.1	2693	9	US-10-115-928-69	Sequence 69, Appl
89	16	12.1	2693	9	US-10-115-928-70	Sequence 70, Appl
90	16	12.1	2817	10	US-09-381-624A-4	Sequence 4, Appl
91	16	12.1	3149	10	US-09-978-168-3	Sequence 3, Appl
92	16	12.1	3437	10	US-09-919-497-34	Sequence 34, Appl

93	16	12.1	3437	10	US-09-954-456-1170	Sequence 1170, Ap	15	11.4	517	10	US-09-924-035A-541	Sequence 541, Ap
94	16	12.1	3705	12	US-10-044-090-307	Sequence 307, App	15	11.4	526	9	US-09-736-457-1473	Sequence 1473, Ap
95	16	12.1	4815	10	US-09-888-615-14	Sequence 14, Appl	15	11.4	526	9	US-09-902-941-1473	Sequence 1473, Ap
96	16	12.1	4863	10	US-09-815-242-4071	Sequence 4071, Ap	15	11.4	526	9	US-09-849-626-1473	Sequence 1473, Ap
97	16	12.1	5184	10	US-09-845-583-9	Sequence 9, Appl	15	11.4	526	9	US-10-017-754-1473	Sequence 1473, Ap
98	16	12.1	6093	9	US-09-764-891-8090	Sequence 8090, Ap	15	11.4	536	10	US-09-864-864-221	Sequence 221, App
99	16	12.1	6640	9	US-10-092-154-1077	Sequence 1077, Ap	15	11.4	674	10	US-09-833-790-141	Sequence 141, App
100	16	12.1	6640	10	US-09-764-847-1077	Sequence 1077, Ap	15	11.4	696	10	US-09-741-669-203	Sequence 203, App
101	16	12.1	6940	10	US-09-880-107-1724	Sequence 1724, Ap	15	11.4	715	9	US-09-815-242-6124	Sequence 6124, Ap
102	16	12.1	7878	10	US-09-962-436-562	Sequence 562, App	15	11.4	782	10	US-09-776-724A-12	Sequence 12, Appl
103	16	12.1	8408	9	US-10-098-841-51	Sequence 51, Appl	15	11.4	838	9	US-09-880-107-3491	Sequence 3491, Ap
104	16	12.1	8420	9	US-10-098-841-50	Sequence 50, Appl	15	11.4	838	9	US-09-989-920-11	Sequence 11, Appl
105	16	12.1	8495	9	US-10-108-605-356	Sequence 356, App	15	11.4	908	9	US-09-764-891-2652	Sequence 2652, App
106	16	12.1	8974	10	US-10-092-154-1078	Sequence 1078, Ap	15	11.4	1032	10	US-09-989-920-12	Sequence 12, Appl
107	16	12.1	8974	10	US-09-764-847-1078	Sequence 1078, Ap	15	11.4	1033	9	US-09-759-468-1	Sequence 1, Appl
108	16	12.1	10923	9	US-10-092-154-1079	Sequence 1079, Ap	15	11.4	1093	9	US-10-103-313-67	Sequence 67, Appl
109	16	12.1	10923	10	US-09-764-847-1079	Sequence 1079, Ap	15	11.4	1132	9	US-09-938-842A-1737	Sequence 1737, Ap
110	16	12.1	12848	9	US-10-114-170-252	Sequence 252, App	15	11.4	1158	9	US-09-808-602-5	Sequence 5, Appl
111	16	12.1	17056	10	US-09-893-238-3	Sequence 3, Appl	15	11.4	1430	9	US-09-800-198-5	Sequence 5, Appl
112	16	12.1	23106	9	US-09-863-049A-1	Sequence 1, Appl	15	11.4	1431	9	US-09-977-418-21	Sequence 21, Appl
113	16	12.1	38155	9	US-10-114-170-79	Sequence 79, Appl	15	11.4	1431	9	US-09-894-633A-90	Sequence 90, Appl
114	16	12.1	48908	9	US-10-114-170-137	Sequence 137, Appl	15	11.4	1477	10	US-09-852-797-12	Sequence 12, Appl
115	16	12.1	65464	9	US-09-859-888-3	Sequence 3, Appl	15	11.4	1526	9	US-09-764-891-180	Sequence 180, App
116	16	12.1	175561	9	US-10-017-721-3	Sequence 3, Appl	15	11.4	1586	9	US-09-852-797-12	Sequence 12, Appl
117	16	12.1	368004	10	US-09-943-634-3	Sequence 3, Appl	15	11.4	1586	10	US-09-853-161-12	Sequence 12, Appl
118	15	11.4	227	9	US-10-066-543-2727	Sequence 2727, Ap	15	11.4	1586	10	US-09-852-797-12	Sequence 12, Appl
119	15	11.4	231	10	US-09-834-975-122	Sequence 122, App	15	11.4	1622	10	US-09-864-864-303	Sequence 303, App
120	15	11.4	242	9	US-10-066-543-1647	Sequence 1647, Ap	15	11.4	1653	9	US-09-938-842A-1054	Sequence 1054, Ap
121	15	11.4	293	9	US-09-440-829-1	Sequence 1, Appl	15	11.4	1674	9	US-09-764-891-8862	Sequence 8862, Ap
122	15	11.4	293	10	US-09-834-975-107	Sequence 107, App	15	11.4	1674	9	US-09-764-891-8865	Sequence 8865, Ap
123	15	11.4	304	9	US-10-025-380-501	Sequence 501, App	15	11.4	1711	9	US-10-114-170-105	Sequence 105, App
124	15	11.4	304	10	US-09-922-217-501	Sequence 501, App	15	11.4	1859	9	US-10-132-744A-5	Sequence 5, Appl
125	15	11.4	304	10	US-09-833-263-501	Sequence 501, App	15	11.4	1859	10	US-09-963-908-2	Sequence 2, Appl
126	15	11.4	347	9	US-10-060-036-1888	Sequence 1888, Ap	15	11.4	1964	10	US-09-963-908-2	Sequence 2, Appl
127	15	11.4	347	9	US-10-060-036-2038	Sequence 2038, Ap	15	11.4	1984	9	US-10-050-704-70	Sequence 70, Appl
128	15	11.4	347	9	US-10-060-036-2040	Sequence 2040, Ap	15	11.4	2000	9	US-09-938-842A-4713	Sequence 4713, Ap
129	15	11.4	347	9	US-10-060-036-3571	Sequence 3571, Ap	15	11.4	2004	10	US-09-887-576-217	Sequence 217, App
130	15	11.4	347	9	US-09-764-891-7699	Sequence 7699, Ap	15	11.4	2012	9	US-09-809-391-132	Sequence 132, App
131	15	11.4	347	9	US-10-066-543-182	Sequence 182, App	15	11.4	2031	10	US-09-764-853-342	Sequence 342, App
132	15	11.4	347	9	US-10-066-543-2353	Sequence 2353, Ap	15	11.4	2053	10	US-09-945-301-118	Sequence 118, App
133	15	11.4	347	9	US-10-066-543-2682	Sequence 2682, Ap	15	11.4	2190	9	US-10-138-927-75	Sequence 75, Appl
134	15	11.4	347	9	US-10-066-543-3123	Sequence 3123, Ap	15	11.4	2215	9	US-10-092-154-1595	Sequence 1595, Ap
135	15	11.4	349	9	US-09-764-891-1334	Sequence 1334, Ap	15	11.4	2215	10	US-09-764-847-1588	Sequence 1588, Ap
136	15	11.4	351	9	US-10-060-036-3572	Sequence 3572, Ap	15	11.4	2219	9	US-10-092-154-1588	Sequence 1588, Ap
137	15	11.4	351	10	US-09-764-853-343	Sequence 343, App	15	11.4	2220	10	US-09-764-847-1588	Sequence 1588, Ap
138	15	11.4	366	10	US-09-834-975-498	Sequence 498, App	15	11.4	2220	10	US-10-092-154-1589	Sequence 1589, Ap
139	15	11.4	374	9	US-09-764-891-542	Sequence 542, App	15	11.4	2593	10	US-09-925-297-62	Sequence 62, Appl
140	15	11.4	386	9	US-09-918-995-35519	Sequence 35519, A	15	11.4	2760	10	US-09-764-864-160	Sequence 160, App
141	15	11.4	389	10	US-09-998-598-2550	Sequence 2550, Ap	15	11.4	3102	9	US-10-121-746-17	Sequence 17, Appl
142	15	11.4	396	9	US-09-970-966-148	Sequence 148, App	15	11.4	3507	10	US-09-834-975-829	Sequence 829, App
143	15	11.4	396	10	US-09-825-294-148	Sequence 148, App	15	11.4	3507	10	US-09-834-975-829	Sequence 829, App
144	15	11.4	417	10	US-09-960-352-933	Sequence 933, App	15	11.4	3594	9	US-09-764-891-9949	Sequence 9949, Ap
145	15	11.4	420	10	US-09-960-352-12483	Sequence 12483, A	15	11.4	3894	9	US-10-028-072-233	Sequence 233, App
146	15	11.4	430	9	US-09-918-995-5960	Sequence 5960, Ap	15	11.4	4333	9	US-10-121-049-233	Sequence 233, App
147	15	11.4	431	10	US-09-800-729-157	Sequence 157, App	15	11.4	4333	9	US-10-123-904-233	Sequence 233, App
148	15	11.4	442	10	US-09-800-729-157	Sequence 157, App	15	11.4	4333	9	US-10-140-470-233	Sequence 233, App
149	15	11.4	443	10	US-09-360-352-13726	Sequence 13726, A	15	11.4	4333	9	US-10-175-746-233	Sequence 233, App
150	15	11.4	445	10	US-09-960-352-13069	Sequence 13069, A	15	11.4	4333	9	US-10-176-918-233	Sequence 233, App
151	15	11.4	446	10	US-09-919-580-640	Sequence 640, App	15	11.4	4333	9	US-10-137-865-233	Sequence 233, App
152	15	11.4	447	9	US-09-933-797-510	Sequence 510, App	15	11.4	4333	9	US-10-142-411-233	Sequence 233, App
153	15	11.4	447	9	US-09-918-995-26903	Sequence 26903, A	15	11.4	4333	9	US-10-143-114-233	Sequence 233, App
154	15	11.4	456	10	US-09-735-705-83	Sequence 83, Appl	15	11.4	4333	9	US-10-140-002-233	Sequence 233, App
155	15	11.4	460	10	US-09-850-716A-83	Sequence 83, Appl	15	11.4	4333	9	US-10-142-419-233	Sequence 233, App
156	15	11.4	460	10	US-09-897-778-83	Sequence 83, Appl	15	11.4	4333	9	US-10-123-262-233	Sequence 233, App
157	15	11.4	461	10	US-09-864-761-6066	Sequence 6066, Ap	15	11.4	4333	9	US-10-141-755-233	Sequence 233, App
158	15	11.4	461	10	US-09-864-761-14441	Sequence 14441, A	15	11.4	4333	9	US-10-141-755-233	Sequence 233, App
159	15	11.4	462	9	US-09-918-995-13118	Sequence 13118, A	15	11.4	4333	9	US-10-143-114-233	Sequence 233, App
160	15	11.4	469	9	US-09-918-995-13890	Sequence 13890, A	15	11.4	4333	9	US-10-140-002-233	Sequence 233, App
161	15	11.4	491	9	US-09-918-995-15899	Sequence 15899, A	15	11.4	4333	9	US-10-142-423-233	Sequence 233, App
162	15	11.4	492	9	US-09-918-995-15899	Sequence 15899, A	15	11.4	4333	9	US-10-142-423-233	Sequence 233, App
163	15	11.4	496	9	US-09-918-995-31964	Sequence 31964, A	15	11.4	4333	9	US-10-141-755-233	Sequence 233, App
164	15	11.4	511	10	US-09-918-995-25194	Sequence 25194, A	15	11.4	4333	9	US-10-141-755-233	Sequence 233, App
165	15	11.4	511	10	US-09-867-701-6033	Sequence 6033, Ap	15	11.4	4333	9	US-10-143-108-233	Sequence 233, App
166	15	11.4	513	9	US-09-918-995-32570	Sequence 32570, A	15	11.4	4333	9	US-10-123-032-233	Sequence 233, App

c 239	15	11.4	4333	9	US-10-123-236-233	Sequence 233, App	c 312	15	11.4	4333	9	US-10-137-868-233	Sequence 233, App
c 240	15	11.4	4333	9	US-10-123-261-233	Sequence 233, App	c 313	15	11.4	4333	9	US-10-146-729-233	Sequence 233, App
c 241	15	11.4	4333	9	US-10-140-921-233	Sequence 233, App	c 314	15	11.4	4333	9	US-10-146-791-233	Sequence 233, App
c 242	15	11.4	4333	9	US-10-140-928-233	Sequence 233, App	c 315	15	11.4	4333	9	US-10-147-484-233	Sequence 233, App
c 243	15	11.4	4333	9	US-10-121-045-233	Sequence 233, App	c 316	15	11.4	4333	9	US-10-147-492-233	Sequence 233, App
c 244	15	11.4	4333	9	US-10-123-239-233	Sequence 233, App	c 317	15	11.4	4333	9	US-10-147-508-233	Sequence 233, App
c 245	15	11.4	4333	9	US-10-123-903-233	Sequence 233, App	c 318	15	11.4	4333	9	US-10-147-512-233	Sequence 233, App
c 246	15	11.4	4333	9	US-10-124-813-233	Sequence 233, App	c 319	15	11.4	4333	9	US-10-158-782-233	Sequence 233, App
c 247	15	11.4	4333	9	US-10-124-822-233	Sequence 233, App	c 320	15	11.4	4333	9	US-10-175-735-233	Sequence 233, App
c 248	15	11.4	4333	9	US-10-140-925-233	Sequence 233, App	c 321	15	11.4	4333	9	US-10-123-905-233	Sequence 233, App
c 249	15	11.4	4333	9	US-10-160-498-233	Sequence 233, App	c 322	15	11.4	4333	9	US-10-123-907-233	Sequence 233, App
c 250	15	11.4	4333	9	US-10-121-041-233	Sequence 233, App	c 323	15	11.4	4333	9	US-10-124-815-233	Sequence 233, App
c 251	15	11.4	4333	9	US-10-121-043-233	Sequence 233, App	c 324	15	11.4	4333	9	US-10-125-921A-233	Sequence 233, App
c 252	15	11.4	4333	9	US-10-121-047-233	Sequence 233, App	c 325	15	11.4	4333	9	US-10-125-928A-233	Sequence 233, App
c 253	15	11.4	4333	9	US-10-123-215-233	Sequence 233, App	c 326	15	11.4	4333	9	US-10-127-821A-233	Sequence 233, App
c 254	15	11.4	4333	9	US-10-123-902-233	Sequence 233, App	c 327	15	11.4	4333	9	US-10-127-822A-233	Sequence 233, App
c 255	15	11.4	4333	9	US-10-123-908-233	Sequence 233, App	c 328	15	11.4	4333	9	US-10-127-824A-233	Sequence 233, App
c 256	15	11.4	4333	9	US-10-123-909-233	Sequence 233, App	c 329	15	11.4	4333	9	US-10-127-826A-233	Sequence 233, App
c 257	15	11.4	4333	9	US-10-123-910-233	Sequence 233, App	c 330	15	11.4	4333	9	US-10-127-827A-233	Sequence 233, App
c 258	15	11.4	4333	9	US-10-124-813-233	Sequence 233, App	c 331	15	11.4	4333	9	US-10-127-828A-233	Sequence 233, App
c 259	15	11.4	4333	9	US-10-124-817-233	Sequence 233, App	c 332	15	11.4	4333	9	US-10-127-830A-233	Sequence 233, App
c 260	15	11.4	4333	9	US-10-124-824-233	Sequence 233, App	c 333	15	11.4	4333	9	US-10-127-832A-233	Sequence 233, App
c 261	15	11.4	4333	9	US-10-125-922-233	Sequence 233, App	c 334	15	11.4	4333	9	US-10-127-833A-233	Sequence 233, App
c 262	15	11.4	4333	9	US-10-125-924-233	Sequence 233, App	c 335	15	11.4	4333	9	US-10-127-834A-233	Sequence 233, App
c 263	15	11.4	4333	9	US-10-127-828A-233	Sequence 233, App	c 336	15	11.4	4333	9	US-10-127-836A-233	Sequence 233, App
c 264	15	11.4	4333	9	US-10-127-829A-233	Sequence 233, App	c 337	15	11.4	4333	9	US-10-127-841A-233	Sequence 233, App
c 265	15	11.4	4333	9	US-10-127-835A-233	Sequence 233, App	c 338	15	11.4	4333	9	US-10-127-844A-233	Sequence 233, App
c 266	15	11.4	4333	9	US-10-127-839A-233	Sequence 233, App	c 339	15	11.4	4333	9	US-10-128-687A-233	Sequence 233, App
c 267	15	11.4	4333	9	US-10-127-901A-233	Sequence 233, App	c 340	15	11.4	4333	9	US-10-128-688A-233	Sequence 233, App
c 268	15	11.4	4333	9	US-10-128-693A-233	Sequence 233, App	c 341	15	11.4	4333	9	US-10-128-689A-233	Sequence 233, App
c 269	15	11.4	4333	9	US-10-131-818A-233	Sequence 233, App	c 342	15	11.4	4333	9	US-10-128-694A-233	Sequence 233, App
c 270	15	11.4	4333	9	US-10-131-822A-233	Sequence 233, App	c 343	15	11.4	4333	9	US-10-131-825A-233	Sequence 233, App
c 271	15	11.4	4333	9	US-10-131-824A-233	Sequence 233, App	c 344	15	11.4	4333	9	US-10-230-417-233	Sequence 233, App
c 272	15	11.4	4333	9	US-10-131-824A-233	Sequence 233, App	c 345	15	11.4	4333	9	US-10-121-051-233	Sequence 233, App
c 273	15	11.4	4333	9	US-10-131-830A-233	Sequence 233, App	c 346	15	11.4	4333	9	US-10-131-815A-233	Sequence 233, App
c 274	15	11.4	4333	9	US-10-131-837A-233	Sequence 233, App	c 347	15	11.4	4333	9	US-10-131-817A-233	Sequence 233, App
c 275	15	11.4	4333	9	US-10-137-872A-233	Sequence 233, App	c 348	15	11.4	4333	9	US-10-131-821A-233	Sequence 233, App
c 276	15	11.4	4333	9	US-10-140-860-233	Sequence 233, App	c 349	15	11.4	4333	9	US-10-131-822A-233	Sequence 233, App
c 277	15	11.4	4333	9	US-10-142-417-233	Sequence 233, App	c 350	15	11.4	4333	9	US-10-131-828A-233	Sequence 233, App
c 278	15	11.4	4333	9	US-10-147-500-233	Sequence 233, App	c 351	15	11.4	4333	9	US-10-131-835A-233	Sequence 233, App
c 279	15	11.4	4333	9	US-10-147-502-233	Sequence 233, App	c 352	15	11.4	4333	9	US-10-137-864A-233	Sequence 233, App
c 280	15	11.4	4333	9	US-10-147-515-233	Sequence 233, App	c 353	15	11.4	4333	9	US-10-137-869A-233	Sequence 233, App
c 281	15	11.4	4333	9	US-10-147-517-233	Sequence 233, App	c 354	15	11.4	4333	9	US-10-147-523-233	Sequence 233, App
c 282	15	11.4	4333	9	US-10-147-519-233	Sequence 233, App	c 355	15	11.4	4333	9	US-10-158-785-233	Sequence 233, App
c 283	15	11.4	4333	9	US-10-147-526-233	Sequence 233, App	c 356	15	11.4	4333	9	US-10-121-042-233	Sequence 233, App
c 284	15	11.4	4333	9	US-10-147-527-233	Sequence 233, App	c 357	15	11.4	4350	9	US-10-125-181-37	Sequence 37, Appl
c 285	15	11.4	4333	9	US-10-152-395-233	Sequence 233, App	c 358	15	11.4	4460	10	US-09-764-864-308	Sequence 308, App
c 286	15	11.4	4333	9	US-10-157-782-233	Sequence 233, App	c 359	15	11.4	4623	10	US-09-864-864-288	Sequence 288, App
c 287	15	11.4	4333	9	US-10-121-040-233	Sequence 233, App	c 360	15	11.4	5080	9	US-10-092-154-1137	Sequence 1137, App
c 288	15	11.4	4333	9	US-10-121-056-233	Sequence 233, App	c 361	15	11.4	5080	9	US-10-092-154-1198	Sequence 1198, App
c 289	15	11.4	4333	9	US-10-121-061-233	Sequence 233, App	c 362	15	11.4	5080	10	US-09-764-847-1197	Sequence 1197, Ap
c 290	15	11.4	4333	9	US-10-123-235-233	Sequence 233, App	c 363	15	11.4	5080	10	US-09-764-847-1198	Sequence 1198, Ap
c 291	15	11.4	4333	9	US-10-124-818-233	Sequence 233, App	c 364	15	11.4	5102	10	US-09-954-456-1870	Sequence 1870, Ap
c 292	15	11.4	4333	9	US-10-125-926A-233	Sequence 233, App	c 365	15	11.4	5289	9	US-10-044-807-1	Sequence 1, Appli
c 293	15	11.4	4333	9	US-10-125-930A-233	Sequence 233, App	c 366	15	11.4	5720	10	US-09-800-729-13	Sequence 18, Appl
c 294	15	11.4	4333	9	US-10-127-831A-233	Sequence 233, App	c 367	15	11.4	5769	10	US-09-764-853-184	Sequence 134, App
c 295	15	11.4	4333	9	US-10-127-837A-233	Sequence 233, App	c 368	15	11.4	7727	9	US-09-764-891-7685	Sequence 7685, Ap
c 296	15	11.4	4333	9	US-10-127-838B-233	Sequence 233, App	c 369	15	11.4	7831	9	US-10-043-487-182	Sequence 182, App
c 297	15	11.4	4333	9	US-10-127-842A-233	Sequence 233, App	c 370	15	11.4	8095	10	US-09-880-107-2228	Sequence 2228, Ap
c 298	15	11.4	4333	9	US-10-127-843A-233	Sequence 233, App	c 371	15	11.4	8689	9	US-09-808-602-78	Sequence 78, Appl
c 299	15	11.4	4333	9	US-10-127-845A-233	Sequence 233, App	c 372	15	11.4	8689	9	US-09-800-198-66	Sequence 66, Appl
c 300	15	11.4	4333	9	US-10-127-846A-233	Sequence 233, App	c 373	15	11.4	9729	9	US-09-808-602-12	Sequence 12, Appl
c 301	15	11.4	4333	9	US-10-127-848A-233	Sequence 233, App	c 374	15	11.4	9729	9	US-09-800-198-12	Sequence 12, Appl
c 302	15	11.4	4333	9	US-10-127-849A-233	Sequence 233, App	c 375	15	11.4	9826	9	US-09-808-602-7	Sequence 7, Appli
c 303	15	11.4	4333	9	US-10-127-850A-233	Sequence 233, App	c 376	15	11.4	9826	9	US-09-800-198-7	Sequence 7, Appli
c 304	15	11.4	4333	9	US-10-127-851A-233	Sequence 233, App	c 377	15	11.4	10319	9	US-09-764-868-1232	Sequence 1232, Ap
c 305	15	11.4	4333	9	US-10-128-684A-233	Sequence 233, App	c 378	15	11.4	15571	10	US-09-880-107-2387	Sequence 2387, Ap
c 306	15	11.4	4333	9	US-10-128-686A-233	Sequence 233, App	c 379	15	11.4	15849	10	US-09-880-107-2362	Sequence 2362, Ap
c 307	15	11.4	4333	9	US-10-128-690A-233	Sequence 233, App	c 380	15	11.4	16428	9	US-10-092-154-1613	Sequence 1613, Ap
c 308	15	11.4	4333	9	US-10-128-691A-233	Sequence 233, App	c 381	15	11.4	16428	10	US-09-764-847-1613	Sequence 1613, Ap
c 309	15	11.4	4333	9	US-10-131-819A-233	Sequence 233, App	c 382	15	11.4	19736	12	US-10-014-502-3	Sequence 3, Appli
c 310	15	11.4	4333	9	US-10-131-828A-233	Sequence 233, App	c 383	15	11.4	20029	9	US-10-092-154-1624	Sequence 1624, Ap
c 311	15	11.4	4333	9	US-10-131-836A-233	Sequence 233, App	c 384	15	11.4	20029	10	US-09-764-847-1624	Sequence 1624, Ap

C 385	15	11.4	31834	9	US-10-092-154-1612	Sequence 1612, Ap	c 458	14	10.6	380	10	US-09-960-352-6073	Sequence 6073, Ap
C 386	15	11.4	31834	10	US-09-764-847-1612	Sequence 1612, Ap	459	14	10.6	381	9	US-09-803-719-2081	Sequence 2081, Ap
C 387	15	11.4	32035	9	US-10-092-154-1611	Sequence 1611, Ap	c 460	14	10.6	385	10	US-09-960-352-928	Sequence 928, App
C 388	15	11.4	32035	10	US-09-764-847-1611	Sequence 1611, Ap	c 461	14	10.6	387	9	US-09-728-444-975	Sequence 975, App
C 389	15	11.4	170834	10	US-09-835-232-7	Sequence 7, Appl	c 462	14	10.6	394	9	US-09-918-995-5037	Sequence 5037, Ap
C 390	15	11.4	197496	9	US-09-877-177-10	Sequence 10, Appl	c 463	14	10.6	397	10	US-09-864-761-5116	Sequence 5116, Ap
C 391	15	11.4	225883	9	US-10-175-523-57	Sequence 57, Appl	c 464	14	10.6	398	10	US-09-960-253-117	Sequence 117, App
C 392	15	11.4	335913	9	US-09-754-853A-2	Sequence 2, Appl	c 465	14	10.6	400	7	US-08-781-986A-4291	Sequence 4291, Ap
C 393	15	11.4	335913	9	US-09-754-853A-3	Sequence 3, Appl	c 466	14	10.6	400	10	US-09-864-761-3628	Sequence 3628, Ap
C 394	15	11.4	659158	9	US-09-771-208-20	Sequence 20, Appl	c 467	14	10.6	400	10	US-09-880-107-1098	Sequence 1098, Ap
C 395	15	11.4	1691139	9	US-10-067-514-1	Sequence 1, Appl	c 468	14	10.6	401	9	US-09-946-807-715	Sequence 715, App
C 396	15	11.4	1691139	9	US-10-067-514-1	Sequence 1, Appl	469	14	10.6	401	9	US-09-946-807-715	Sequence 715, App
C 397	14	10.6	27	10	US-09-828-313-128	Sequence 128, App	470	14	10.6	401	10	US-09-795-668-715	Sequence 715, App
C 398	14	10.6	30	9	US-09-973-025-34	Sequence 34, Appl	471	14	10.6	401	10	US-09-795-668-716	Sequence 716, App
C 399	14	10.6	30	9	US-09-899-303-34	Sequence 34, Appl	472	14	10.6	401	10	US-09-795-668-715	Sequence 715, App
C 400	14	10.6	30	9	US-09-995-808-34	Sequence 34, Appl	473	14	10.6	401	10	US-09-795-668-716	Sequence 716, App
C 401	14	10.6	30	9	US-09-780-533A-3325	Sequence 3325, Ap	474	14	10.6	401	10	US-09-918-995-5037	Sequence 5037, Ap
C 402	14	10.6	38	9	US-09-780-164-1290	Sequence 1290, Ap	475	14	10.6	403	9	US-09-918-995-5037	Sequence 5037, Ap
C 403	14	10.6	81	10	US-09-864-761-21838	Sequence 21838, A	c 476	14	10.6	404	10	US-09-960-352-13218	Sequence 13218, A
C 404	14	10.6	122	10	US-09-864-761-21838	Sequence 21838, A	c 477	14	10.6	410	10	US-09-960-352-3698	Sequence 3698, Ap
C 405	14	10.6	137	10	US-09-864-761-21838	Sequence 21838, A	c 478	14	10.6	410	10	US-09-960-352-6586	Sequence 6586, Ap
C 406	14	10.6	140	10	US-09-864-761-21838	Sequence 21838, A	c 479	14	10.6	412	9	US-09-918-995-3096	Sequence 3096, Ap
C 407	14	10.6	151	10	US-09-764-891-66	Sequence 66, Appl	c 480	14	10.6	418	9	US-09-918-995-32996	Sequence 32996, A
C 408	14	10.6	157	10	US-09-864-761-24892	Sequence 24892, A	c 481	14	10.6	418	9	US-09-918-995-32996	Sequence 32996, A
C 409	14	10.6	158	10	US-09-923-876-3467	Sequence 3467, Ap	c 482	14	10.6	422	9	US-09-918-995-3954	Sequence 3954, Ap
C 410	14	10.6	172	9	US-10-060-036-2954	Sequence 2954, Ap	c 483	14	10.6	423	10	US-09-960-352-9386	Sequence 9386, Ap
C 411	14	10.6	175	10	US-09-867-701-1232	Sequence 1232, Ap	c 484	14	10.6	424	10	US-09-998-598-519	Sequence 519, App
C 412	14	10.6	191	10	US-09-864-761-26396	Sequence 26396, A	c 485	14	10.6	426	9	US-09-918-995-36190	Sequence 36190, A
C 413	14	10.6	193	10	US-09-864-761-26396	Sequence 26396, A	c 486	14	10.6	429	9	US-09-918-995-36190	Sequence 36190, A
C 414	14	10.6	198	10	US-09-864-761-21770	Sequence 21770, A	c 487	14	10.6	429	10	US-09-864-761-4259	Sequence 4259, A
C 415	14	10.6	198	10	US-09-864-761-25452	Sequence 25452, A	c 488	14	10.6	429	10	US-09-560-863-938	Sequence 938, App
C 416	14	10.6	211	10	US-09-864-761-30924	Sequence 30924, A	c 489	14	10.6	429	10	US-09-983-965-2539	Sequence 2539, Ap
C 417	14	10.6	211	10	US-09-960-352-3006	Sequence 3006, Ap	c 490	14	10.6	430	10	US-09-960-352-172	Sequence 172, App
C 418	14	10.6	213	8	US-08-796-570A-6	Sequence 6, Appl	c 491	14	10.6	432	9	US-09-918-995-36063	Sequence 36063, A
C 419	14	10.6	213	10	US-10-067-477-6	Sequence 6, Appl	c 492	14	10.6	432	9	US-09-918-995-36063	Sequence 36063, A
C 420	14	10.6	216	10	US-09-864-761-24166	Sequence 24166, A	c 493	14	10.6	445	9	US-09-918-995-26477	Sequence 26477, A
C 421	14	10.6	235	10	US-09-867-701-6943	Sequence 6943, A	c 494	14	10.6	445	10	US-09-783-590-3731	Sequence 3731, Ap
C 422	14	10.6	250	10	US-09-924-035A-491	Sequence 491, App	c 495	14	10.6	448	9	US-09-918-995-29023	Sequence 29023, A
C 423	14	10.6	251	9	US-09-535-459-1899	Sequence 1899, Ap	c 496	14	10.6	448	9	US-09-918-995-29023	Sequence 29023, A
C 424	14	10.6	265	10	US-09-878-574-10977	Sequence 10977, A	c 497	14	10.6	463	9	US-09-918-995-28383	Sequence 28383, A
C 425	14	10.6	270	10	US-10-066-543-2633	Sequence 2633, Ap	c 498	14	10.6	464	9	US-09-918-995-27489	Sequence 27489, A
C 426	14	10.6	280	10	US-09-878-574-14797	Sequence 14797, A	c 499	14	10.6	466	9	US-09-918-995-11266	Sequence 11266, A
C 427	14	10.6	282	9	US-09-938-842A-4480	Sequence 4480, Ap	c 500	14	10.6	468	10	US-09-864-761-619	Sequence 619, App
C 428	14	10.6	285	10	US-09-887-576-707	Sequence 707, App	c 501	14	10.6	468	9	US-09-918-995-13248	Sequence 13248, A
C 429	14	10.6	290	9	US-09-535-459-1931	Sequence 1931, Ap	c 502	14	10.6	471	10	US-09-864-761-1293	Sequence 1293, Ap
C 430	14	10.6	290	10	US-09-294-093B-3243	Sequence 3243, Ap	c 503	14	10.6	472	10	US-09-918-995-11303	Sequence 11303, A
C 431	14	10.6	293	10	US-09-878-574-14704	Sequence 14704, A	c 504	14	10.6	472	10	US-09-864-761-2616	Sequence 2616, Ap
C 432	14	10.6	295	10	US-09-294-093B-4928	Sequence 4928, Ap	c 505	14	10.6	473	9	US-10-152-661-355	Sequence 355, App
C 433	14	10.6	311	10	US-09-783-590-1846	Sequence 1846, Ap	c 506	14	10.6	473	9	US-09-866-050A-355	Sequence 355, App
C 434	14	10.6	321	9	US-09-864-761-19338	Sequence 19338, A	c 507	14	10.6	473	10	US-09-918-995-27832	Sequence 27832, A
C 435	14	10.6	331	9	US-09-738-626-3336	Sequence 3336, Ap	c 508	14	10.6	474	10	US-09-867-701-5248	Sequence 5248, Ap
C 436	14	10.6	331	9	US-10-070-676-15	Sequence 15, Appl	c 509	14	10.6	475	9	US-09-918-995-32592	Sequence 32592, A
C 437	14	10.6	339	9	US-09-535-459-1263	Sequence 1263, Ap	c 510	14	10.6	476	10	US-09-864-761-5167	Sequence 5167, Ap
C 438	14	10.6	343	10	US-09-294-093B-107	Sequence 107, App	c 511	14	10.6	477	10	US-09-918-995-27851	Sequence 27851, Ap
C 439	14	10.6	347	10	US-09-834-975-296	Sequence 296, App	c 512	14	10.6	477	10	US-09-917-800A-1242	Sequence 1242, Ap
C 440	14	10.6	352	10	US-09-878-574-5303	Sequence 5303, Ap	c 513	14	10.6	479	9	US-09-918-995-9663	Sequence 9663, Ap
C 441	14	10.6	356	9	US-10-046-935-1053	Sequence 1053, Ap	c 514	14	10.6	486	10	US-09-918-995-27300	Sequence 27300, A
C 442	14	10.6	356	9	US-10-046-935-1244	Sequence 1244, Ap	c 515	14	10.6	486	10	US-09-864-761-1360	Sequence 1360, Ap
C 443	14	10.6	356	9	US-10-042-125A-13	Sequence 13, Appl	c 516	14	10.6	488	9	US-09-918-995-29138	Sequence 29138, A
C 444	14	10.6	356	9	US-09-878-178-1053	Sequence 1053, Ap	c 517	14	10.6	488	9	US-09-918-995-1650	Sequence 1650, A
C 445	14	10.6	356	9	US-09-878-178-1244	Sequence 1244, Ap	c 518	14	10.6	491	9	US-09-918-995-6573	Sequence 6573, Ap
C 446	14	10.6	356	9	US-09-878-178-1480	Sequence 1480, Ap	c 519	14	10.6	494	9	US-09-918-995-1361	Sequence 1361, A
C 447	14	10.6	356	9	US-10-146-502-1053	Sequence 1053, Ap	c 520	14	10.6	496	10	US-09-833-381-1545	Sequence 1545, Ap
C 448	14	10.6	356	9	US-10-146-502-1244	Sequence 1244, Ap	c 521	14	10.6	497	10	US-09-878-574-4839	Sequence 4839, Ap
C 449	14	10.6	356	9	US-10-146-502-1480	Sequence 1480, Ap	c 522	14	10.6	499	9	US-09-918-995-38029	Sequence 38029, Ap
C 450	14	10.6	356	9	US-10-066-543-2582	Sequence 2582, Ap	c 523	14	10.6	500	10	US-09-991-936-1762	Sequence 1762, Ap
C 451	14	10.6	357	10	US-09-867-701-4366	Sequence 4366, Ap	c 524	14	10.6	500	10	US-09-864-761-7753	Sequence 7753, Ap
C 452	14	10.6	359	10	US-09-878-574-3339	Sequence 3339, A	c 525	14	10.6	501	9	US-09-918-995-1764	Sequence 1764, Ap
C 453	14	10.6	370	9	US-09-918-995-37583	Sequence 37583, A	c 526	14	10.6	505	9	US-09-918-995-20056	Sequence 20056, A
C 454	14	10.6	372	9	US-10-152-661-60	Sequence 60, Appl	c 527	14	10.6	506	9	US-09-918-995-31662	Sequence 31662, A
C 455	14	10.6	372	9	US-09-866-050A-60	Sequence 60, Appl	c 528	14	10.6	510	9	US-09-918-995-11644	Sequence 11644, A
C 456	14	10.6	373	9	US-09-918-995-37603	Sequence 37603, A	c 529	14	10.6	510	10	US-09-878-574-1324	Sequence 1324, Ap
C 457	14	10.6	379	10	US-09-960-352-164	Sequence 164, App	c 530	14	10.6	516	9	US-09-978-295A-484	Sequence 484, App

c 531	14	10.6	9	US-09-978-697-484	Sequence 484, App	c 604	14	10.6	10	US-09-960-253-106	Sequence 106, App
c 532	14	10.6	9	US-09-978-192A-484	Sequence 484, App	605	14	10.6	10	US-09-833-381-1457	Sequence 1457, App
c 533	14	10.6	9	US-09-999-832A-484	Sequence 484, App	606	14	10.6	10	US-09-773-774A-3	Sequence 3, Appli
c 534	14	10.6	9	US-09-978-189-484	Sequence 484, App	607	14	10.6	10	US-09-771-961-3	Sequence 3, Appli
c 535	14	10.6	9	US-09-978-608A-484	Sequence 484, App	608	14	10.6	10	US-09-815-242-9391	Sequence 9391, Ap
c 536	14	10.6	9	US-09-978-191A-484	Sequence 484, App	609	14	10.6	10	US-09-815-242-9515	Sequence 9515, Ap
c 537	14	10.6	9	US-09-978-403A-484	Sequence 484, App	c 610	14	10.6	9	US-09-738-626-2788	Sequence 2788, Ap
c 538	14	10.6	9	US-09-978-564A-484	Sequence 484, App	611	14	10.6	10	US-09-815-242-6918	Sequence 6918, Ap
c 539	14	10.6	9	US-09-978-585A-484	Sequence 484, App	c 612	14	10.6	10	US-09-746-491-55	Sequence 55, Appl
c 540	14	10.6	9	US-10-017-081A-484	Sequence 484, App	613	14	10.6	876	US-09-815-242-9175	Sequence 9175, Ap
c 541	14	10.6	9	US-09-978-824-484	Sequence 484, App	c 614	14	10.6	884	US-09-770-445-516	Sequence 516, App
c 542	14	10.6	9	US-09-981-915A-484	Sequence 484, App	c 615	14	10.6	896	US-09-938-842A-3702	Sequence 3702, Ap
c 543	14	10.6	9	US-09-999-833A-484	Sequence 484, App	616	14	10.6	918	US-09-974-300-4772	Sequence 4772, Ap
c 544	14	10.6	9	US-10-167-749-484	Sequence 484, App	617	14	10.6	919	US-09-925-297-14	Sequence 14, Appl
c 545	14	10.6	9	US-09-918-585A-484	Sequence 484, App	618	14	10.6	924	US-09-816-714-3	Sequence 3, Appli
c 546	14	10.6	9	US-09-978-423A-484	Sequence 484, App	619	14	10.6	948	US-09-815-242-7678	Sequence 7678, Ap
c 547	14	10.6	9	US-10-013-921A-484	Sequence 484, App	620	14	10.6	954	US-09-738-396-3	Sequence 3, Appli
c 548	14	10.6	9	US-09-978-193A-484	Sequence 484, App	621	14	10.6	984	US-09-773-774A-1	Sequence 1, Appli
c 549	14	10.6	9	US-10-013-929A-484	Sequence 484, App	622	14	10.6	984	US-09-771-961-1	Sequence 1, Appli
c 550	14	10.6	9	US-10-016-177A-484	Sequence 484, App	623	14	10.6	1033	US-10-228-794-73	Sequence 73, Appl
c 551	14	10.6	9	US-09-999-830A-484	Sequence 484, App	c 624	14	10.6	1089	US-10-139-483-1	Sequence 1, Appli
c 552	14	10.6	9	US-09-978-757A-484	Sequence 484, App	625	14	10.6	1098	US-09-741-669-160	Sequence 160, App
c 553	14	10.6	9	US-09-978-187B-484	Sequence 484, App	626	14	10.6	1099	US-09-263-959-245	Sequence 245, App
c 554	14	10.6	9	US-09-864-761-15327	Sequence 15327, A	c 627	14	10.6	1113	US-09-796-965-1	Sequence 1, Appli
c 555	14	10.6	10	US-09-864-761-12828	Sequence 12828, A	628	14	10.6	1142	US-09-263-959-248	Sequence 248, App
c 556	14	10.6	9	US-09-833-790-349	Sequence 349, App	629	14	10.6	1151	US-09-974-879-105	Sequence 105, App
c 557	14	10.6	9	US-09-854-133-642	Sequence 642, App	c 630	14	10.6	1151	US-09-305-736-105	Sequence 105, App
c 558	14	10.6	9	US-09-833-790-312	Sequence 312, App	c 631	14	10.6	1160	US-09-764-891-5481	Sequence 5481, Ap
c 559	14	10.6	9	US-09-998-598-1386	Sequence 1386, Ap	c 632	14	10.6	1161	US-09-770-445-64	Sequence 64, Appl
c 560	14	10.6	9	US-09-907-969-138	Sequence 138, App	633	14	10.6	1179	US-09-738-396-1	Sequence 1, Appli
c 561	14	10.6	9	US-09-864-761-7394	Sequence 7394, App	634	14	10.6	1185	US-09-938-842A-2520	Sequence 2520, Ap
c 562	14	10.6	9	US-09-884-441-138	Sequence 138, App	635	14	10.6	1208	US-09-925-302-10	Sequence 10, Appl
c 563	14	10.6	9	US-09-918-995-30660	Sequence 30660, A	c 636	14	10.6	1216	US-09-746-491-21	Sequence 21, Appl
c 564	14	10.6	9	US-09-918-995-27673	Sequence 27673, A	637	14	10.6	1228	US-09-263-959-246	Sequence 246, App
c 565	14	10.6	9	US-09-864-761-16140	Sequence 16140, A	c 638	14	10.6	1272	US-08-796-570A-2	Sequence 2, Appli
c 566	14	10.6	9	US-10-092-154-1940	Sequence 1940, Ap	c 639	14	10.6	1272	US-10-067-477-2	Sequence 2, Appli
c 567	14	10.6	9	US-09-764-847-1940	Sequence 1940, Ap	640	14	10.6	1279	US-09-974-879-68	Sequence 68, Appl
c 568	14	10.6	9	US-10-092-154-471	Sequence 471, App	641	14	10.6	1286	US-09-305-736-68	Sequence 68, Appl
c 569	14	10.6	9	US-09-764-847-471	Sequence 471, App	c 642	14	10.6	1286	US-09-746-491-56	Sequence 56, Appl
c 570	14	10.6	9	US-09-764-868-531	Sequence 531, App	c 643	14	10.6	1286	US-09-746-491-57	Sequence 57, Appl
c 571	14	10.6	9	US-09-535-459-1930	Sequence 1930, Ap	c 644	14	10.6	1293	US-09-945-527-36	Sequence 36, Appl
c 572	14	10.6	9	US-09-918-995-29374	Sequence 29374, A	645	14	10.6	1300	US-09-263-959-251	Sequence 251, App
c 573	14	10.6	9	US-09-864-761-12272	Sequence 12272, A	646	14	10.6	1311	US-09-738-626-3441	Sequence 3441, Ap
c 574	14	10.6	9	US-09-864-761-7661	Sequence 7661, Ap	647	14	10.6	1321	US-10-174-590-371	Sequence 371, App
c 575	14	10.6	9	US-09-535-459-1928	Sequence 1928, Ap	648	14	10.6	1321	US-10-176-758-371	Sequence 371, App
c 576	14	10.6	9	US-10-178-782-1	Sequence 1, Appli	649	14	10.6	1321	US-10-175-737-371	Sequence 371, App
c 577	14	10.6	9	US-09-864-761-9702	Sequence 9702, Ap	650	14	10.6	1321	US-10-173-706-371	Sequence 371, App
c 578	14	10.6	9	US-09-834-975-421	Sequence 421, App	651	14	10.6	1321	US-10-175-738-371	Sequence 371, App
c 579	14	10.6	9	US-10-056-543-2782	Sequence 2782, Ap	652	14	10.6	1321	US-10-175-752-371	Sequence 371, App
c 580	14	10.6	9	US-09-864-761-8150	Sequence 8150, Ap	653	14	10.6	1321	US-10-176-482-371	Sequence 371, App
c 581	14	10.6	9	US-09-535-459-1929	Sequence 1929, Ap	654	14	10.6	1321	US-10-176-757-371	Sequence 371, App
c 582	14	10.6	9	US-09-864-761-6845	Sequence 6845, Ap	655	14	10.6	1321	US-10-176-913-371	Sequence 371, App
c 583	14	10.6	9	US-09-907-969-106	Sequence 106, App	656	14	10.6	1321	US-10-180-552-371	Sequence 371, App
c 584	14	10.6	9	US-09-884-441-106	Sequence 106, App	657	14	10.6	1321	US-10-180-557-371	Sequence 371, App
c 585	14	10.6	9	US-09-864-761-8793	Sequence 8793, Ap	658	14	10.6	1321	US-10-173-700-371	Sequence 371, App
c 586	14	10.6	9	US-09-864-761-9068	Sequence 9068, Ap	659	14	10.6	1321	US-10-174-572-371	Sequence 371, App
c 587	14	10.6	9	US-09-864-761-9075	Sequence 9075, Ap	660	14	10.6	1321	US-10-174-579-371	Sequence 371, App
c 588	14	10.6	9	US-09-864-761-12817	Sequence 12817, A	661	14	10.6	1321	US-10-174-582-371	Sequence 371, App
c 589	14	10.6	9	US-09-764-891-1170	Sequence 1170, Ap	662	14	10.6	1321	US-10-174-588-371	Sequence 371, App
c 590	14	10.6	9	US-09-925-301-439	Sequence 439, App	663	14	10.6	1321	US-10-175-739-371	Sequence 371, App
c 591	14	10.6	9	US-09-834-975-597	Sequence 597, App	664	14	10.6	1321	US-10-175-740-371	Sequence 371, App
c 592	14	10.6	9	US-09-770-149-966	Sequence 966, App	665	14	10.6	1321	US-10-175-743-371	Sequence 371, App
c 593	14	10.6	9	US-09-917-800A-252	Sequence 252, App	666	14	10.6	1321	US-10-176-488-371	Sequence 371, App
c 594	14	10.6	9	US-09-998-598-463	Sequence 463, Appl	667	14	10.6	1321	US-10-176-492-371	Sequence 371, App
c 595	14	10.6	9	US-09-910-664-23	Sequence 23, Appli	668	14	10.6	1321	US-10-176-747-371	Sequence 371, App
c 596	14	10.6	9	US-09-917-800A-1280	Sequence 1280, Ap	669	14	10.6	1321	US-10-176-750-371	Sequence 371, App
c 597	14	10.6	9	US-09-764-891-1349	Sequence 1349, Ap	670	14	10.6	1321	US-10-176-985-371	Sequence 371, App
c 598	14	10.6	9	US-09-834-975-357	Sequence 357, App	671	14	10.6	1321	US-10-176-987-371	Sequence 371, App
c 599	14	10.6	9	US-09-747-155-115	Sequence 115, App	672	14	10.6	1321	US-10-176-991-371	Sequence 371, App
c 600	14	10.6	9	US-09-960-253-107	Sequence 107, App	673	14	10.6	1321	US-10-176-992-371	Sequence 371, App
c 601	14	10.6	9	US-09-764-877-2132	Sequence 2132, Ap	674	14	10.6	1321	US-10-176-993-371	Sequence 371, App
c 602	14	10.6	9	US-09-925-300-330	Sequence 330, App	675	14	10.6	1321	US-10-184-658-371	Sequence 371, App
c 603	14	10.6	12	US-10-033-528-1830	Sequence 1830, App	676	14	10.6	1321	US-10-173-695-371	Sequence 371, App


```
969 14 10.6 1321 9 US-10-205-898-371 Sequence 371, App
970 14 10.6 1321 9 US-10-205-901-371 Sequence 371, App
971 14 10.6 1321 9 US-10-205-903-371 Sequence 371, App
972 14 10.6 1321 9 US-10-206-909-371 Sequence 371, App
973 14 10.6 1321 9 US-10-206-910-371 Sequence 371, App
974 14 10.6 1321 9 US-10-206-911-371 Sequence 371, App
975 14 10.6 1321 9 US-10-206-912-371 Sequence 371, App
976 14 10.6 1321 9 US-10-206-913-371 Sequence 371, App
977 14 10.6 1321 9 US-10-206-914-371 Sequence 371, App
978 14 10.6 1321 9 US-10-206-920-371 Sequence 371, App
979 14 10.6 1321 9 US-10-206-921-371 Sequence 371, App
980 14 10.6 1321 9 US-10-206-923-371 Sequence 371, App
981 14 10.6 1321 9 US-10-206-925-371 Sequence 371, App
982 14 10.6 1321 9 US-10-206-926-371 Sequence 371, App
983 14 10.6 1321 9 US-10-206-927-371 Sequence 371, App
984 14 10.6 1321 9 US-10-207-916-371 Sequence 371, App
985 14 10.6 1321 9 US-10-207-917-371 Sequence 371, App
986 14 10.6 1321 9 US-10-207-918-371 Sequence 371, App
987 14 10.6 1321 9 US-10-207-919-371 Sequence 371, App
988 14 10.6 1321 9 US-10-207-920-371 Sequence 371, App
989 14 10.6 1321 9 US-10-207-923-371 Sequence 371, App
990 14 10.6 1321 9 US-10-208-021-371 Sequence 371, App
991 14 10.6 1321 9 US-10-208-022-371 Sequence 371, App
992 14 10.6 1321 9 US-10-208-023-371 Sequence 371, App
993 14 10.6 1321 9 US-10-208-026-371 Sequence 371, App
994 14 10.6 1321 9 US-10-208-029-371 Sequence 371, App
995 14 10.6 1321 9 US-10-208-030-371 Sequence 371, App
996 14 10.6 1321 9 US-10-232-232-371 Sequence 371, App
997 14 10.6 1321 9 US-09-946-374-365 Sequence 365, App
998 14 10.6 1321 9 US-10-012-121A-365 Sequence 365, App
999 14 10.6 1321 9 US-10-015-869A-365 Sequence 371, App
1000 14 10.6 1321 12 US-10-052-586-371
```

ALIGNMENTS

```
RESULT 1
US-10-160-293-3
; Sequence 3, Application US/10160293
; Publication No. US20030022208A1
; GENERAL INFORMATION:
; APPLICANT: LINQUIST, Erika et al.
; TITLE OF INVENTION: ISOLATED HUMAN SECRETED PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN SECRETED PROTEINS, AND
; TITLE OF INVENTION: USES THEREOF
; FILE REFERENCE: CL001241-PROV
; CURRENT APPLICATION NUMBER: US/10/160,293
; CURRENT FILING DATE: 2002-06-04
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 39776
; TYPE: DNA
; ORGANISM: Human
US-10-160-293-3
```

```
Query Match 14.4%; Score 19; DB 9; Length 39776;
Best Local Similarity 100.0%; Pred. No. 1.3;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 57 GGCCACAGAGATGCGAAG 75
Db 18251 GGCCACAGAGATGCGAAG 18269
|||||
```

```
RESULT 2
US-09-919-580-517
; Sequence 517, Application US/09919580
; Patent No. US20020110832A1
; GENERAL INFORMATION:
; APPLICANT: Pyle, Ruth
; APPLICANT: Xu, Jiangchun
```

```
; APPLICANT: Secrist, Heather
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.552
; CURRENT APPLICATION NUMBER: US/09/919,580
; CURRENT FILING DATE: 2001-07-30
; NUMBER OF SEQ ID NOS: 934
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 517
; LENGTH: 243
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 243
; OTHER INFORMATION: n = A,T,C or G
US-09-919-580-517
```

```
Query Match 13.6%; Score 18; DB 10; Length 243;
Best Local Similarity 100.0%; Pred. No. 5.4;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 40 GATAAGGCCCAAGCTGAAG 57
|||||
```

```
Db 102 GATAAGGCCCAAGCTGAAG 119
|||||
```

RESULT 3

```
US-09-919-580-224
; Sequence 224, Application US/09919580
; Patent No. US20020110832A1
; GENERAL INFORMATION:
; APPLICANT: Pyle, Ruth
; APPLICANT: Xu, Jiangchun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF COLON CANCER
; FILE REFERENCE: 210121.552
; CURRENT APPLICATION NUMBER: US/09/919,580
; CURRENT FILING DATE: 2001-07-30
; NUMBER OF SEQ ID NOS: 934
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 224
; LENGTH: 353
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-919-580-224
```

```
Query Match 13.6%; Score 18; DB 10; Length 353;
Best Local Similarity 100.0%; Pred. No. 5.3;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 40 GATAAGGCCCAAGCTGAAG 57
|||||
```

```
Db 76 GATAAGGCCCAAGCTGAAG 93
|||||
```

RESULT 4

```
US-09-919-580-207
; Sequence 207, Application US/09919580
; Patent No. US20020110832A1
; GENERAL INFORMATION:
; APPLICANT: Pyle, Ruth
; APPLICANT: Xu, Jiangchun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.552
; CURRENT APPLICATION NUMBER: US/09/919,580
; CURRENT FILING DATE: 2001-07-30
; NUMBER OF SEQ ID NOS: 934
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 207
```

; LENGTH: 395
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-919-580-207

Query Match 13.6%; Score 18; DB 10; Length 395;
Best Local Similarity 100.0%; Pred. No. 5.3;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 40 GATAGGCCAAGCTGAAG 57
|||||

Db 111 GATAGGCCAAGCTGAAG 128
|||||

RESULT 5

US-09-960-352-10885
; Sequence 10885, Application US/09960352
; Patent No. US20020137139A1

; GENERAL INFORMATION:

; APPLICANT: Warren, Wesley C.

; APPLICANT: Byatt, John C.

; APPLICANT: Mathialagan, Nagappan

; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
; FILE OF INVENTION: MUSCLE AND FAT DEPOSITION

; FILE REFERENCE: 16511.006/37-21(10298)C

; CURRENT APPLICATION NUMBER: US/09/960,352

; CURRENT FILING DATE: 2001-09-24

; NUMBER OF SEQ ID NOS: 15112

; SEQ ID NO 10885

; LENGTH: 404

; TYPE: DNA

; ORGANISM: Bos taurus

; OTHER INFORMATION: Clone ID: 47-LIB188-013-Q1-E1-D4

US-09-960-352-10885

Query Match 13.6%; Score 18; DB 10; Length 404;
Best Local Similarity 100.0%; Pred. No. 5.3;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 40 GATAGGCCAAGCTGAAG 57
|||||

Db 153 GATAGGCCAAGCTGAAG 170
|||||

RESULT 6

US-09-960-352-6906
; Sequence 6906, Application US/09960352
; Patent No. US20020137139A1

; GENERAL INFORMATION:

; APPLICANT: Warren, Wesley C.

; APPLICANT: Byatt, John C.

; APPLICANT: Mathialagan, Nagappan

; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
; FILE OF INVENTION: MUSCLE AND FAT DEPOSITION

; FILE REFERENCE: 16511.006/37-21(10298)C

; CURRENT APPLICATION NUMBER: US/09/960,352

; CURRENT FILING DATE: 2001-09-24

; NUMBER OF SEQ ID NOS: 15112

; SEQ ID NO 6906

; LENGTH: 405

; TYPE: DNA

; ORGANISM: Bos taurus

; OTHER INFORMATION: Clone ID: 30-LIB188-021-Q1-E1-H5

US-09-960-352-6906

Query Match 13.6%; Score 18; DB 10; Length 405;
Best Local Similarity 100.0%; Pred. No. 5.3;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 40 GATAGGCCAAGCTGAAG 57
|||||

Db 153 GATAGGCCAAGCTGAAG 170
|||||

RESULT 7

US-09-960-352-12808
; Sequence 12808, Application US/09960352
; Patent No. US20020137139A1

; GENERAL INFORMATION:

; APPLICANT: Warren, Wesley C.

; APPLICANT: Byatt, John C.

; APPLICANT: Mathialagan, Nagappan

; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION
; FILE OF INVENTION: MUSCLE AND FAT DEPOSITION

; FILE REFERENCE: 16511.006/37-21(10298)C

; CURRENT APPLICATION NUMBER: US/09/960,352

; CURRENT FILING DATE: 2001-09-24

; NUMBER OF SEQ ID NOS: 15112

; SEQ ID NO 12808

; LENGTH: 408

; TYPE: DNA

; ORGANISM: Bos taurus

; OTHER INFORMATION: Clone ID: 55-LIB188-007-Q1-E1-F12

US-09-960-352-12808

Query Match 13.6%; Score 18; DB 10; Length 408;
Best Local Similarity 100.0%; Pred. No. 5.3;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 40 GATAGGCCAAGCTGAAG 57
|||||

Db 164 GATAGGCCAAGCTGAAG 181
|||||

RESULT 8

US-09-998-598-71
; Sequence 71, Application US/09998598
; Patent No. US20020150922A1

; GENERAL INFORMATION:

; APPLICANT: Stolk, John A.

; APPLICANT: Xu, Jiangchun

; APPLICANT: Chenault, Ruth A.

; APPLICANT: Meagher, Madelein Joy

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE OF INVENTION: DIAGNOSIS OF COLON CANCER

; FILE REFERENCE: 210121.561

; CURRENT APPLICATION NUMBER: US/09/998,598

; CURRENT FILING DATE: 2001-11-16

; NUMBER OF SEQ ID NOS: 2606

; SOFTWARE: Corixa Invention Disclosure Database

; SEQ ID NO 71

; LENGTH: 426

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-998-598-71

Query Match 13.6%; Score 18; DB 10; Length 426;
Best Local Similarity 100.0%; Pred. No. 5.3;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 40 GATAGGCCAAGCTGAAG 57
|||||

Db 87 GATAGGCCAAGCTGAAG 104
|||||

RESULT 9

US-09-919-580-503
; Sequence 503, Application US/09919580
; Patent No. US20020110832A1

; GENERAL INFORMATION:

; APPLICANT: Pyle, Ruth

; APPLICANT: Xu, Jiangchun

; APPLICANT: Secrist, Heather

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.552
; CURRENT APPLICATION NUMBER: US/09/919,580
; CURRENT FILING DATE: 2001-07-30
; NUMBER OF SEQ ID NOS: 934
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 503
; LENGTH: 428
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-919-580-503

Query Match 13.6%; Score 18; DB 10; Length 428;
Best Local Similarity 100.0%; Pred. No. 5.3;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 40 GATAAGGCCCAAGCTGAAG 57
|||||
Db 89 GATAAGGCCCAAGCTGAAG 106

RESULT 10
US-09-919-580-633
; Sequence 633, Application US/09919580
; Patent No. US20020110832A1
; GENERAL INFORMATION:
; APPLICANT: Pyle, Ruth
; APPLICANT: Xu, Jiangchun
; APPLICANT: Secrist, Heather
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.552
; CURRENT APPLICATION NUMBER: US/09/919,580
; CURRENT FILING DATE: 2001-07-30
; NUMBER OF SEQ ID NOS: 934
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 633
; LENGTH: 429
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-919-580-633

Query Match 13.6%; Score 18; DB 10; Length 429;
Best Local Similarity 100.0%; Pred. No. 5.3;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 40 GATAAGGCCCAAGCTGAAG 57
|||||
Db 90 GATAAGGCCCAAGCTGAAG 107

RESULT 11
US-09-919-580-627
; Sequence 627, Application US/09919580
; Patent No. US20020110832A1
; GENERAL INFORMATION:
; APPLICANT: Pyle, Ruth
; APPLICANT: Xu, Jiangchun
; APPLICANT: Secrist, Heather
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.552
; CURRENT APPLICATION NUMBER: US/09/919,580
; CURRENT FILING DATE: 2001-07-30
; NUMBER OF SEQ ID NOS: 934
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 627
; LENGTH: 434
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-919-580-627

Query Match 13.6%; Score 18; DB 10; Length 434;
Best Local Similarity 100.0%; Pred. No. 5.3;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 40 GATAAGGCCCAAGCTGAAG 57
|||||
Db 99 GATAAGGCCCAAGCTGAAG 116

RESULT 12
US-09-919-580-458
; Sequence 458, Application US/09919580
; Patent No. US20020110832A1
; GENERAL INFORMATION:
; APPLICANT: Pyle, Ruth
; APPLICANT: Xu, Jiangchun
; APPLICANT: Secrist, Heather
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.552
; CURRENT APPLICATION NUMBER: US/09/919,580
; CURRENT FILING DATE: 2001-07-30
; NUMBER OF SEQ ID NOS: 934
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 458
; LENGTH: 438
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-919-580-458

Query Match 13.6%; Score 18; DB 10; Length 438;
Best Local Similarity 100.0%; Pred. No. 5.3;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 40 GATAAGGCCCAAGCTGAAG 57
|||||
Db 99 GATAAGGCCCAAGCTGAAG 116

RESULT 13
US-09-919-580-583
; Sequence 583, Application US/09919580
; Patent No. US20020110832A1
; GENERAL INFORMATION:
; APPLICANT: Pyle, Ruth
; APPLICANT: Xu, Jiangchun
; APPLICANT: Secrist, Heather
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.552
; CURRENT APPLICATION NUMBER: US/09/919,580
; CURRENT FILING DATE: 2001-07-30
; NUMBER OF SEQ ID NOS: 934
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 583
; LENGTH: 438
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-919-580-583

Query Match 13.6%; Score 18; DB 10; Length 438;
Best Local Similarity 100.0%; Pred. No. 5.3;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 40 GATAAGGCCCAAGCTGAAG 57
|||||
Db 99 GATAAGGCCCAAGCTGAAG 116

RESULT 14
US-09-919-580-910
; Sequence 910, Application US/09919580
; Patent No. US20020110832A1

; GENERAL INFORMATION:
; APPLICANT: Pyle, Ruth
; APPLICANT: Xu, Jiangchun
; APPLICANT: Secretist, Heather
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF COLON CANCER
; FILE REFERENCE: 210121.552
; CURRENT FILING DATE: 2001-07-30
; NUMBER OF SEQ ID NOS: 934
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 910
; LENGTH: 439
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: 1, 438, 439
; OTHER INFORMATION: n = A,T,C or G

US-09-919-580-910

Query Match 13.6%; Score 18; DB 10; Length 439;
Best Local Similarity 100.0%; Pred. No. 5.3;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 40 GATAAGGCCAAGCTGAAG 57
|||||
Db 94 GATAAGGCCAAGCTGAAG 111

RESULT 15

US-09-919-580-659
; Sequence 659, Application US/09919580
; Patent No. US20020110832A1

; GENERAL INFORMATION:
; APPLICANT: Pyle, Ruth
; APPLICANT: Xu, Jiangchun
; APPLICANT: Secretist, Heather
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF COLON CANCER
; FILE REFERENCE: 210121.552
; CURRENT FILING DATE: 2001-07-30
; NUMBER OF SEQ ID NOS: 934
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 659
; LENGTH: 443
; TYPE: DNA
; ORGANISM: Homo sapiens

US-09-919-580-659

Query Match 13.6%; Score 18; DB 10; Length 443;
Best Local Similarity 100.0%; Pred. No. 5.3;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 40 GATAAGGCCAAGCTGAAG 57
|||||
Db 104 GATAAGGCCAAGCTGAAG 121

RESULT 16

US-09-919-580-727
; Sequence 727, Application US/09919580
; Patent No. US20020110832A1

; GENERAL INFORMATION:
; APPLICANT: Pyle, Ruth
; APPLICANT: Xu, Jiangchun
; APPLICANT: Secretist, Heather
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF COLON CANCER
; FILE REFERENCE: 210121.552
; CURRENT FILING DATE: 2001-07-30

; NUMBER OF SEQ ID NOS: 934
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 727
; LENGTH: 445
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-919-580-727

Query Match 13.6%; Score 18; DB 10; Length 445;
Best Local Similarity 100.0%; Pred. No. 5.2;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 40 GATAAGGCCAAGCTGAAG 57
|||||
Db 99 GATAAGGCCAAGCTGAAG 116

RESULT 17

US-09-919-580-891
; Sequence 891, Application US/09919580
; Patent No. US20020110832A1

; GENERAL INFORMATION:
; APPLICANT: Pyle, Ruth
; APPLICANT: Xu, Jiangchun
; APPLICANT: Secretist, Heather
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF COLON CANCER
; FILE REFERENCE: 210121.552
; CURRENT FILING DATE: 2001-07-30
; NUMBER OF SEQ ID NOS: 934
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 891
; LENGTH: 446
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: 427, 444
; OTHER INFORMATION: n = A,T,C or G

US-09-919-580-891

Query Match 13.6%; Score 18; DB 10; Length 446;
Best Local Similarity 100.0%; Pred. No. 5.2;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 40 GATAAGGCCAAGCTGAAG 57
|||||
Db 83 GATAAGGCCAAGCTGAAG 100

RESULT 18

US-09-880-107-3158/c
; Sequence 3158, Application US/09880107
; Patent No. US20020142981A1

; GENERAL INFORMATION:
; APPLICANT: Horne, Darci T.
; APPLICANT: Vockley, Joseph G.
; APPLICANT: Scherf, Uwe
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
; FILE REFERENCE: 44921-5028-WO
; CURRENT FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 60/211,379
; PRIOR FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: US 60/237,054
; PRIOR FILING DATE: 2000-10-02
; NUMBER OF SEQ ID NOS: 3950
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3158
; LENGTH: 446
; TYPE: DNA

```

; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20020142981A1 T59161
; NAME/KEY: unsure
; LOCATION: (1)...(446)
; OTHER INFORMATION: n = a or c or g or t
US-09-880-107-3158

Query Match      13.6%; Score 18; DB 10; Length 446;
Best Local Similarity 100.0%; Pred. No. 5.2;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      40 GATAAGGCCCAAGCTGAAG 57
      |||||||
Db      336 GATAAGGCCCAAGCTGAAG 319

RESULT 19
US-09-918-995-15768
; Sequence 13768, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; PRIOR FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 15768
; LENGTH: 449
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-918-995-15768

Query Match      13.6%; Score 18; DB 9; Length 449;
Best Local Similarity 100.0%; Pred. No. 5.2;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      40 GATAAGGCCCAAGCTGAAG 57
      |||||||
Db      146 GATAAGGCCCAAGCTGAAG 163

RESULT 20
US-09-919-580-122
; Sequence 122, Application US/09919580
; Patent No. US20020110832A1
; GENERAL INFORMATION:
; APPLICANT: Pyle, Ruth
; APPLICANT: Xu, Jiangchun
; APPLICANT: Secrist, Heather
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.552
; CURRENT APPLICATION NUMBER: US/09/919,580
; CURRENT FILING DATE: 2001-07-30
; NUMBER OF SEQ ID NOS: 934
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 122
; LENGTH: 449
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-919-580-122

Query Match      13.6%; Score 18; DB 10; Length 449;
Best Local Similarity 100.0%; Pred. No. 5.2;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      40 GATAAGGCCCAAGCTGAAG 57

```

```

      |||||||
Db      112 GATAAGGCCCAAGCTGAAG 129

RESULT 21
US-09-919-580-190
; Sequence 190, Application US/09919580
; Patent No. US20020110832A1
; GENERAL INFORMATION:
; APPLICANT: Pyle, Ruth
; APPLICANT: Xu, Jiangchun
; APPLICANT: Secrist, Heather
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.552
; CURRENT APPLICATION NUMBER: US/09/919,580
; CURRENT FILING DATE: 2001-07-30
; NUMBER OF SEQ ID NOS: 934
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 190
; LENGTH: 451
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-919-580-190

Query Match      13.6%; Score 18; DB 10; Length 451;
Best Local Similarity 100.0%; Pred. No. 5.2;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      40 GATAAGGCCCAAGCTGAAG 57
      |||||||
Db      112 GATAAGGCCCAAGCTGAAG 129

RESULT 22
US-09-919-580-553
; Sequence 553, Application US/09919580
; Patent No. US20020110832A1
; GENERAL INFORMATION:
; APPLICANT: Pyle, Ruth
; APPLICANT: Xu, Jiangchun
; APPLICANT: Secrist, Heather
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.552
; CURRENT APPLICATION NUMBER: US/09/919,580
; CURRENT FILING DATE: 2001-07-30
; NUMBER OF SEQ ID NOS: 934
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 553
; LENGTH: 451
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-919-580-553

Query Match      13.6%; Score 18; DB 10; Length 451;
Best Local Similarity 100.0%; Pred. No. 5.2;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      40 GATAAGGCCCAAGCTGAAG 57
      |||||||
Db      112 GATAAGGCCCAAGCTGAAG 129

RESULT 23
US-09-919-580-348
; Sequence 348, Application US/09919580
; Patent No. US20020110832A1
; GENERAL INFORMATION:
; APPLICANT: Pyle, Ruth
; APPLICANT: Xu, Jiangchun
; APPLICANT: Secrist, Heather
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY

```

; TITLE OF INVENTION: AND DIAGNOSIS OF COLON CANCER
; FILE REFERENCE: 210121.552
; CURRENT APPLICATION NUMBER: US/09/919,580
; CURRENT FILING DATE: 2001-07-30
; NUMBER OF SEQ ID NOS: 934
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 348
; LENGTH: 452
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-919-580-348

Query Match 13.6%; Score 18; DB 10; Length 452;
Best Local Similarity 100.0%; Pred. No. 5.2;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 40 GATAAGGCCAAGCTGAAG 57
|||||
Db 113 GATAAGGCCAAGCTGAAG 130

RESULT 24

US-09-919-580-274
; Sequence 274, Application US/09919580
; Patent No. US20020110832A1
; GENERAL INFORMATION:
; APPLICANT: Pyle, Ruth
; APPLICANT: Xu, Jiangchun
; APPLICANT: Secrist, Heather
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF COLON CANCER
; FILE REFERENCE: 210121.552
; CURRENT APPLICATION NUMBER: US/09/919,580
; CURRENT FILING DATE: 2001-07-30
; NUMBER OF SEQ ID NOS: 934
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 274
; LENGTH: 453
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-919-580-274

Query Match 13.6%; Score 18; DB 10; Length 453;
Best Local Similarity 100.0%; Pred. No. 5.2;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 40 GATAAGGCCAAGCTGAAG 57
|||||
Db 114 GATAAGGCCAAGCTGAAG 131

RESULT 25

US-09-880-107-3019
; Sequence 3019, Application US/09880107
; Patent No. US20020142981A1
; GENERAL INFORMATION:
; APPLICANT: Horne, Darci T.
; APPLICANT: Vockley, Joseph G.
; APPLICANT: Scherf, Uwe
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
; FILE REFERENCE: 44921-5028-WO
; CURRENT APPLICATION NUMBER: US/09/880,107
; CURRENT FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 60/211,379
; PRIOR FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: US 60/237,054
; PRIOR FILING DATE: 2000-10-02
; SOFTWARE: PatentIn Ver. 2.1
; NUMBER OF SEQ ID NOS: 3950
; SEQ ID NO 3019
; LENGTH: 453
; TYPE: DNA

; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20020142981A1 S54005
US-09-880-107-3019

Query Match 13.6%; Score 18; DB 10; Length 453;
Best Local Similarity 100.0%; Pred. No. 5.2;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 40 GATAAGGCCAAGCTGAAG 57
|||||
Db 105 GATAAGGCCAAGCTGAAG 122

RESULT 26

US-09-919-580-273
; Sequence 273, Application US/09919580
; Patent No. US20020110832A1
; GENERAL INFORMATION:
; APPLICANT: Pyle, Ruth
; APPLICANT: Xu, Jiangchun
; APPLICANT: Secrist, Heather
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF COLON CANCER
; FILE REFERENCE: 210121.552
; CURRENT APPLICATION NUMBER: US/09/919,580
; CURRENT FILING DATE: 2001-07-30
; NUMBER OF SEQ ID NOS: 934
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 273
; LENGTH: 455
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-919-580-273

Query Match 13.6%; Score 18; DB 10; Length 455;
Best Local Similarity 100.0%; Pred. No. 5.2;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 40 GATAAGGCCAAGCTGAAG 57
|||||
Db 112 GATAAGGCCAAGCTGAAG 129

RESULT 27

US-09-919-580-219
; Sequence 219, Application US/09919580
; Patent No. US20020110832A1
; GENERAL INFORMATION:
; APPLICANT: Pyle, Ruth
; APPLICANT: Xu, Jiangchun
; APPLICANT: Secrist, Heather
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF COLON CANCER
; FILE REFERENCE: 210121.552
; CURRENT APPLICATION NUMBER: US/09/919,580
; CURRENT FILING DATE: 2001-07-30
; NUMBER OF SEQ ID NOS: 934
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 219
; LENGTH: 456
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-919-580-219

Query Match 13.6%; Score 18; DB 10; Length 456;
Best Local Similarity 100.0%; Pred. No. 5.2;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 40 GATAAGGCCAAGCTGAAG 57
|||||
Db 117 GATAAGGCCAAGCTGAAG 134

```
RESULT 28
US-09-960-352-2605/c
; Sequence 2605, Application US/09960352
; Patent No. US20020137139A1
; GENERAL INFORMATION:
; APPLICANT: Warten, Wesley C.
; APPLICANT: Tao, Nengbing
; APPLICANT: Byatt, John C.
; APPLICANT: Mathialagan, Nagappan
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
; FILE REFERENCE: 16511.006/37-21(10298)C
; CURRENT APPLICATION NUMBER: US/09/960.352
; CURRENT FILING DATE: 2001-09-24
; NUMBER OF SEQ ID NOS: 15112
; SEQ ID NO 2605
; LENGTH: 461
; TYPE: DNA
; ORGANISM: Bos taurus
; NAME/KEY: unsure
; LOCATION: (278)
; OTHER INFORMATION: unsure at all n locations
; OTHER INFORMATION: Clone ID: 12-LIB188-023-Q1-El-C7
US-09-960-352-2605

Query Match      13.6%; Score 18; DB 10; Length 461;
Best Local Similarity 100.0%; Pred. No. 5.2;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 40 GATAAGGCCCAAGCTGAAG 57
    |||||||
Db 403 GATAAGGCCCAAGCTGAAG 386

RESULT 29
US-09-919-580-56
; Sequence 56, Application US/09919580
; Patent No. US20020110832A1
; GENERAL INFORMATION:
; APPLICANT: Pyle, Ruth
; APPLICANT: Xu, Jiangchun
; APPLICANT: Secrist, Heather
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.552
; CURRENT APPLICATION NUMBER: US/09/919,580
; CURRENT FILING DATE: 2001-07-30
; NUMBER OF SEQ ID NOS: 934
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 56
; LENGTH: 465
; TYPE: DNA
; ORGANISM: Homo sapiens
; NAME/KEY: misc.feature
; LOCATION: 404_446, 459
; OTHER INFORMATION: n = A,T,C or G
US-09-919-580-56

Query Match      13.6%; Score 18; DB 10; Length 465;
Best Local Similarity 100.0%; Pred. No. 5.2;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 40 GATAAGGCCCAAGCTGAAG 57
    |||||||
Db 101 GATAAGGCCCAAGCTGAAG 118

RESULT 30
US-09-918-995-21161
; Sequence 21161, Application US/09918995
```

```
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 21161
; LENGTH: 468
; TYPE: DNA
; ORGANISM: Homo sapiens
; NAME/KEY: misc.feature
; LOCATION: (1)...(468)
; OTHER INFORMATION: n = A,T,C or G
US-09-918-995-21161

Query Match      13.6%; Score 18; DB 9; Length 468;
Best Local Similarity 100.0%; Pred. No. 5.2;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 40 GATAAGGCCCAAGCTGAAG 57
    |||||||
Db 118 GATAAGGCCCAAGCTGAAG 135

RESULT 31
US-09-918-995-24062
; Sequence 24062, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 24062
; LENGTH: 469
; TYPE: DNA
; ORGANISM: Homo sapiens
; NAME/KEY: misc.feature
; LOCATION: (1)...(469)
; OTHER INFORMATION: n = A,T,C or G
US-09-918-995-24062

Query Match      13.6%; Score 18; DB 9; Length 469;
Best Local Similarity 100.0%; Pred. No. 5.2;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 40 GATAAGGCCCAAGCTGAAG 57
    |||||||
Db 108 GATAAGGCCCAAGCTGAAG 125

RESULT 32
US-09-918-995-15273
; Sequence 15273, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
```

FILE REFERENCE: 20411-756
CURRENT APPLICATION NUMBER: US/09/918,995
CURRENT FILING DATE: 2001-07-30
PRIOR APPLICATION NUMBER: US/09/235,076
PRIOR FILING DATE: 1999-01-20
NUMBER OF SEQ ID NOS: 38054
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 15273
LENGTH: 473
TYPE: DNA
ORGANISM: Homo sapiens
NAME/KEY: misc_feature
LOCATION: (1)...(473)
OTHER INFORMATION: n = A,T,C or G
US-09-918-995-15273

Query Match 13.6%; Score 18; DB 9; Length 473;
Best Local Similarity 100.0%; Pred. No. 5.2;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 40 GATAGGCCCAAGCTGAAG 57
|||||
DB 142 GATAGGCCCAAGCTGAAG 159

RESULT 33

US-09-918-995-22477
Sequence 22477, Application US/09918995
Publication No. US20030073623A1
GENERAL INFORMATION:
APPLICANT: Hyseq, Inc.
TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
FILE REFERENCE: 20411-756
CURRENT APPLICATION NUMBER: US/09/918,995
CURRENT FILING DATE: 2001-07-30
PRIOR APPLICATION NUMBER: US/09/235,076
PRIOR FILING DATE: 1999-01-20
NUMBER OF SEQ ID NOS: 38054
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 22477
LENGTH: 474
TYPE: DNA
ORGANISM: Homo sapiens
NAME/KEY: misc_feature
LOCATION: (1)...(474)
OTHER INFORMATION: n = A,T,C or G
US-09-918-995-22477

Query Match 13.6%; Score 18; DB 9; Length 474;
Best Local Similarity 100.0%; Pred. No. 5.2;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 40 GATAGGCCCAAGCTGAAG 57
|||||
DB 68 GATAGGCCCAAGCTGAAG 85

RESULT 34

US-09-918-995-19906
Sequence 19906, Application US/09918995
Publication No. US20030073623A1
GENERAL INFORMATION:
APPLICANT: Hyseq, Inc.
TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
FILE REFERENCE: 20411-756
CURRENT APPLICATION NUMBER: US/09/918,995
CURRENT FILING DATE: 2001-07-30
PRIOR APPLICATION NUMBER: US/09/235,076
PRIOR FILING DATE: 1999-01-20

NUMBER OF SEQ ID NOS: 38054
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 19906
LENGTH: 483
TYPE: DNA
ORGANISM: Homo sapiens
NAME/KEY: misc_feature
LOCATION: (1)...(483)
OTHER INFORMATION: n = A,T,C or G
US-09-918-995-19906

Query Match 13.6%; Score 18; DB 9; Length 483;
Best Local Similarity 100.0%; Pred. No. 5.2;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 40 GATAGGCCCAAGCTGAAG 57
|||||
DB 130 GATAGGCCCAAGCTGAAG 147

RESULT 35

US-09-918-995-24178/c
Sequence 24178, Application US/09918995
Publication No. US20030073623A1
GENERAL INFORMATION:
APPLICANT: Hyseq, Inc.
TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
FILE REFERENCE: 20411-756
CURRENT APPLICATION NUMBER: US/09/918,995
CURRENT FILING DATE: 2001-07-30
PRIOR APPLICATION NUMBER: US/09/235,076
PRIOR FILING DATE: 1999-01-20
NUMBER OF SEQ ID NOS: 38054
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 24178
LENGTH: 494
TYPE: DNA
ORGANISM: Homo sapiens
NAME/KEY: misc_feature
LOCATION: (1)...(494)
OTHER INFORMATION: n = A,T,C or G
US-09-918-995-24178

Query Match 13.6%; Score 18; DB 9; Length 494;
Best Local Similarity 100.0%; Pred. No. 5.2;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 40 GATAGGCCCAAGCTGAAG 57
|||||
DB 384 GATAGGCCCAAGCTGAAG 367

RESULT 36

US-09-919-580-874
Sequence 874, Application US/09919580
Patent No. US20020110832A1
GENERAL INFORMATION:
APPLICANT: Pyle, Ruth
APPLICANT: Xu, Jiangchun
APPLICANT: Secrist, Heather
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
TITLE OF INVENTION: AND DIAGNOSIS OF COLON CANCER
FILE REFERENCE: 210121.552
CURRENT APPLICATION NUMBER: US/09/919,580
CURRENT FILING DATE: 2001-07-30
NUMBER OF SEQ ID NOS: 934
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 874
LENGTH: 517
TYPE: DNA

```
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 446, 503, 505
; OTHER INFORMATION: n = A,T,C or G
US-09-919-580-874

Query Match      13.6%; Score 18; DB 10; Length 517;
Best Local Similarity 100.0%; Pred. No. 5.2;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      40 GATAAGGCCAAGCTGAAG 57
      |||
Db      103 GATAAGGCCAAGCTGAAG 120

RESULT 37
US-09-919-580-159
; Sequence 159, Application US/09919580
; Patent No. US20020110832A1
; GENERAL INFORMATION:
; APPLICANT: Pyle, Ruth
; APPLICANT: Xu, Jiangchun
; TITLE OF INVENTION: SECRET, Heather
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.552
; CURRENT APPLICATION NUMBER: US/09/919,580
; CURRENT FILING DATE: 2001-07-30
; NUMBER OF SEQ ID NOS: 934
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 159
; LENGTH: 559
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 414, 556
; OTHER INFORMATION: n = A,T,C or G
US-09-919-580-159

Query Match      13.6%; Score 18; DB 10; Length 559;
Best Local Similarity 100.0%; Pred. No. 5.2;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      40 GATAAGGCCAAGCTGAAG 57
      |||
Db      90 GATAAGGCCAAGCTGAAG 107

RESULT 38
US-09-919-580-798
; Sequence 798, Application US/09919580
; Patent No. US20020110832A1
; GENERAL INFORMATION:
; APPLICANT: Pyle, Ruth
; APPLICANT: Xu, Jiangchun
; TITLE OF INVENTION: SECRET, Heather
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.552
; CURRENT APPLICATION NUMBER: US/09/919,580
; CURRENT FILING DATE: 2001-07-30
; NUMBER OF SEQ ID NOS: 934
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 798
; LENGTH: 693
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 428, 429, 431, 481, 487, 490, 492, 494, 497, 505, 511, 516,
; LOCATION: 539, 554, 562, 575, 597, 641, 664, 682

; OTHER INFORMATION: n = A,T,C or G
US-09-919-580-798

Query Match      13.6%; Score 18; DB 10; Length 693;
Best Local Similarity 100.0%; Pred. No. 5.2;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      40 GATAAGGCCAAGCTGAAG 57
      |||
Db      80 GATAAGGCCAAGCTGAAG 97

RESULT 39
US-10-092-154-1336
; Sequence 1336, Application US/10092154
; Publication No. US20030054375A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC009C1
; CURRENT APPLICATION NUMBER: US/10/092,154
; CURRENT FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 2003
; Prior Application removed - See File Wrapper or Palm
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1336
; LENGTH: 9953
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-092-154-1336

Query Match      13.6%; Score 18; DB 9; Length 9953;
Best Local Similarity 100.0%; Pred. No. 4.6;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      43 AAGGCCAAGCTGAAGGCC 60
      |||
Db      1721 AAGGCCAAGCTGAAGGCC 1738

RESULT 40
US-09-764-847-1336
; Sequence 1336, Application US/09764847
; Patent No. US20020132767A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC009
; CURRENT APPLICATION NUMBER: US/09/764,847
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 2003
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1336
; LENGTH: 9953
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-764-847-1336

Query Match      13.6%; Score 18; DB 10; Length 9953;
Best Local Similarity 100.0%; Pred. No. 4.6;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      43 AAGGCCAAGCTGAAGGCC 60
      |||
Db      1721 AAGGCCAAGCTGAAGGCC 1738

RESULT 41
US-10-092-154-1335
; Sequence 1335, Application US/10092154
; Publication No. US20030054375A1
; GENERAL INFORMATION:
```

```
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC009C1
; CURRENT APPLICATION NUMBER: US/10/092,154
; CURRENT FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 2003
; Prior Application removed - See File Wrapper or Palm
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1335
; LENGTH: 11189
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-092-154-1335

Query Match      13.6%; Score 18; DB 9; Length 11189;
Best Local Similarity 100.0%; Pred. No. 4.6;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      43 AAGGCCAAGCTGAAGGCC 60
      |||
Db      2957 AAGGCCAAGCTGAAGGCC 2974

RESULT 42
US-09-764-847-1335
; Sequence 1335, Application US/09764847
; Patent No. US20020132767A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC009
; CURRENT APPLICATION NUMBER: US/09/764,847
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 2003
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1335
; LENGTH: 11189
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-847-1335

Query Match      13.6%; Score 18; DB 10; Length 11189;
Best Local Similarity 100.0%; Pred. No. 4.6;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      43 AAGGCCAAGCTGAAGGCC 60
      |||
Db      2957 AAGGCCAAGCTGAAGGCC 2974

RESULT 43
US-09-918-995-6168/c
; Sequence 6168, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 6168
; LENGTH: 425
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-918-995-6168

Query Match      12.9%; Score 17; DB 9; Length 425;
```

```
Best Local Similarity 100.0%; Pred. No. 18;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      50 AGCTGAAGGCCACAGAG 66
      |||
Db      104 AGCTGAAGGCCACAGAG 88

RESULT 44
US-09-960-352-5941/c
; Sequence 5941, Application US/09960352
; Patent No. US20020137139A1
; GENERAL INFORMATION:
; APPLICANT: Warren, Wesley C.
; APPLICANT: Tao, Nengbing
; APPLICANT: Byatt, John C.
; APPLICANT: Mathalagan, Nagappan
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION
; TITLE OF INVENTION: MUSCLE AND FAT DEPOSITION
; FILE REFERENCE: 16511.006/37-21(10298)C
; CURRENT APPLICATION NUMBER: US/09/960,352
; CURRENT FILING DATE: 2001-09-24
; NUMBER OF SEQ ID NOS: 15112
; SEQ ID NO 5941
; LENGTH: 428
; TYPE: DNA
; ORGANISM: Bos taurus
; OTHER INFORMATION: Clone ID: 26-BOVMS1-020-Q1-EI-G5
US-09-960-352-5941

Query Match      12.9%; Score 17; DB 10; Length 428;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      75 GAACACTCTGATGACCA 91
      |||
Db      51 GAACACTCTGATGACCA 35

RESULT 45
US-09-918-995-4250/c
; Sequence 4250, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 4250
; LENGTH: 433
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(433)
; OTHER INFORMATION: n = A,T,C or G
US-09-918-995-4250

Query Match      12.9%; Score 17; DB 9; Length 433;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      50 AGCTGAAGGCCACAGAG 66
      |||
Db      104 AGCTGAAGGCCACAGAG 88
```

RESULT 46

US-09-876-889-308/c
; Sequence 308, Application US/09876889
; Patent No. US20020076715A1
; GENERAL INFORMATION:
; APPLICANT: Benson, Darin R.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: King, Gordon E.

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR OVARIAN
; TITLE OF INVENTION: CANCER THERAPY AND DIAGNOSIS

; FILE REFERENCE: 210121.466C3
; CURRENT APPLICATION NUMBER: US/09/876,889
; CURRENT FILING DATE: 2001-06-06

; NUMBER OF SEQ ID NOS: 353
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 308

; LENGTH: 439
; TYPE: DNA

; ORGANISM: Homo sapiens
US-09-876-889-308

Query Match 12.9%; Score 17; DB 10; Length 439;
Best Local Similarity 100.0%; Pred. No. 18;

Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 50 AGCTGAAGGCCACAGAG 66

Db 401 AGCTGAAGGCCACAGAG 385

RESULT 47

US-09-918-995-22450
; Sequence 22450, Application US/09918995
; Publication No. US20030073623A1

; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.

; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; FILE REFERENCE: 20411-756

; CURRENT APPLICATION NUMBER: US/09/918,995
; CURRENT FILING DATE: 2001-07-30

; PRIOR APPLICATION NUMBER: US/09/235,076
; PRIOR FILING DATE: 1999-01-20

; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 22450

; LENGTH: 461
; TYPE: DNA

; ORGANISM: Homo sapiens
US-09-918-995-22450

; NAME/KEY: misc_feature
; LOCATION: (1)...(461)

; OTHER INFORMATION: n = A,T,C or G

Query Match 12.9%; Score 17; DB 9; Length 461;
Best Local Similarity 100.0%; Pred. No. 18;

Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 40 GATAAGGCCACAGCTGAA 56

Db 131 GATAAGGCCACAGCTGAA 147

RESULT 48

US-09-918-995-23067
; Sequence 23067, Application US/09918995
; Publication No. US20030073623A1

; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.

; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; FILE REFERENCE: 20411-756

; CURRENT APPLICATION NUMBER: US/09/918,995
; CURRENT FILING DATE: 2001-07-30

; PRIOR APPLICATION NUMBER: US/09/235,076
; PRIOR FILING DATE: 1999-01-20

; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 22450

; LENGTH: 461
; TYPE: DNA

; FILE REFERENCE: 20411-756

; CURRENT APPLICATION NUMBER: US/09/918,995
; CURRENT FILING DATE: 2001-07-30

; PRIOR APPLICATION NUMBER: US/09/235,076
; PRIOR FILING DATE: 1999-01-20

; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 23067

; LENGTH: 465
; TYPE: DNA

; ORGANISM: Homo sapiens
US-09-918-995-23067

; NAME/KEY: misc_feature
; LOCATION: (1)...(465)

; OTHER INFORMATION: n = A,T,C or G

Query Match 12.9%; Score 17; DB 9; Length 465;
Best Local Similarity 100.0%; Pred. No. 18;

Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 40 GATAAGGCCACAGCTGAA 56

Db 162 GATAAGGCCACAGCTGAA 178

RESULT 49

US-10-102-806-146/c

; Sequence 146, Application US/10102806
; Publication No. US20030054421A1

; GENERAL INFORMATION:
; APPLICANT: Rosen et al.

; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: P103P1C1

; CURRENT APPLICATION NUMBER: US/10/102,806
; CURRENT FILING DATE: 2002-03-22

; PRIOR APPLICATION NUMBER: 09/925,298
; PRIOR FILING DATE: 2001-08-10

; PRIOR APPLICATION NUMBER: PCT/US00/05881
; PRIOR FILING DATE: 2000-03-08

; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12

; NUMBER OF SEQ ID NOS: 846
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 146

; LENGTH: 814
; TYPE: DNA

; ORGANISM: Homo sapiens
US-10-102-806-146

; NAME/KEY: misc_feature
; LOCATION: (1)...(814)

; OTHER INFORMATION: n = A,T,C or G

Query Match 12.9%; Score 17; DB 9; Length 814;
Best Local Similarity 100.0%; Pred. No. 17;

Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 50 AGCTGAAGGCCACAGAG 66

Db 688 AGCTGAAGGCCACAGAG 672

RESULT 50

US-09-738-626-1039
; Sequence 1039, Application US/09738626
; Publication No. US20020197605A1

; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI

; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: P103P1C1

; CURRENT APPLICATION NUMBER: US/10/102,806
; CURRENT FILING DATE: 2002-03-22

; PRIOR APPLICATION NUMBER: 09/925,298
; PRIOR FILING DATE: 2001-08-10

; PRIOR APPLICATION NUMBER: PCT/US00/05881
; PRIOR FILING DATE: 2000-03-08

; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12

; NUMBER OF SEQ ID NOS: 846
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 146

; LENGTH: 814
; TYPE: DNA

; ORGANISM: Homo sapiens
US-10-102-806-146

; NAME/KEY: misc_feature
; LOCATION: (1)...(814)

; OTHER INFORMATION: n = A,T,C or G

; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 1039
; LENGTH: 1077
; TYPE: DNA
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-1039

Query Match 12.98; Score 17; DB 9; Length 1077;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 62 CACAGATGCAGAGAAC 78
|||||
Db 690 CACAGATGCAGAGAAC 706

Search completed: June 2, 2003, 17:18:33
Job time : 135 secs

THIS PAGE BLANK (1/3/79)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 2, 2003, 14:37:52 ; Search time 214 Seconds
(without alignments)
1389.083 Million cell updates/sec

Title: US-09-915-178-1

Perfect score: 132

Sequence: 1 atggcacacaactagacct.....agtggagtgaatttctctga 132

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 2185239 seqs, 1125999159 residues

Word size : 10

Total number of hits satisfying chosen parameters: 258815

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database : N_Geneseq_101002.*

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.*	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.*	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.*	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.*	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.*	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.*	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.*	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.*	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.*	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.*	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.*	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.*	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.*	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.*	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.*	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.*	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.*	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.*	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.*	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.*	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.*	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.*	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.*	

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	132	100.0	147	24	ABA92649
2	132	100.0	147	24	ABA92650
3	132	100.0	147	24	ABA92651
4	18	13.6	65	24	ABN29874
5	18	13.6	203	23	ABV08667
6	18	13.6	243	24	ABK55047
7	18	13.6	283	21	AAC00909
8	18	13.6	353	24	ABK54754
9	18	13.6	395	24	ABK54737

Human prostate exp	399	13.6	23	ABV38568
Human colon cancer	428	13.6	24	ABK55033
Human colon cancer	429	13.6	24	ABK55163
Human colon cancer	434	13.6	24	ABK55157
Human colon cancer	438	13.6	24	ABK54988
Human colon cancer	438	13.6	24	ABK55113
Human colon cancer	439	13.6	24	ABK55440
Human colon cancer	443	13.6	24	ABK55189
Human colon cancer	445	13.6	24	ABK55257
Gene #3159 used to	446	13.6	24	ABN96661
Human colon cancer	446	13.6	24	ABK55421
Human colon cancer	449	13.6	24	ABK54652
Human colon cancer	451	13.6	24	ABK54720
Human colon cancer	451	13.6	24	ABK55083
Human colon cancer	452	13.6	24	ABK54878
Human secreted pro	453	13.6	21	RAC10525
Gene #3020 used to	453	13.6	24	ABN96522
Human colon cancer	453	13.6	24	ABK54804
Human colon cancer	455	13.6	24	ABK54803
Human colon cancer	456	13.6	24	ABK54749
Human colon cancer	465	13.6	24	ABK54586
Human secreted pro	488	13.6	21	RAC10524
Human thymosin b-1	491	13.6	22	RAD03640
Human colon cancer	517	13.6	24	ABK55404
Human prostate exp	534	13.6	23	ABV33875
Human prostate exp	534	13.6	23	ABV42758
Human colon cancer	559	13.6	24	ABK54689
Human colon cancer	693	13.6	24	ABK55328
Human prostate exp	736	13.6	23	ABV25565
Human prostate exp	792	13.6	23	ABV03577
Human prostate exp	836	13.6	23	ABV12746
Human prostate exp	1201	13.6	23	ABV21155
Human prostate exp	1201	13.6	23	ABV22339
Human prostate exp	1201	13.6	23	ABV28163
Human cDNA differe	1262	13.6	22	ABK84119
DNA encoding molec	1517	13.6	22	RAA02039
Drosophila melanog	2606	13.6	23	ABL06285
Drosophila melanog	2609	13.6	23	ABL21125
Human polynucleoti	3291	13.6	22	AAK52974
Human polynucleoti	3754	13.6	22	AAK51990
Drosophila melanog	4229	13.6	23	ABL25848
Drosophila melanog	5032	13.6	23	ABL21124
Drosophila melanog	5401	13.6	23	ABL06284
Genomic sequence #	9953	13.6	23	ABK42449
Human ABC1 genomic	11189	13.6	23	ABK42448
Human spliced tran	183999	12.9	22	AAF92831
Human gene signatu	60	12.9	24	ABN41936
Human immune/haema	139	12.9	16	AAK26247
Human immune/haema	261	12.9	22	AAK79934
Human immune/haema	261	12.9	22	AAK79937
Human cDNA for nov	261	12.9	22	AAK79938
CDNA encoding thym	428	12.9	24	ABK34292
Thymosin-beta-10-1	430	12.9	22	AAK00689
zea mays DNA fragm	430	12.9	24	AAK38157
DNA encoding novel	490	12.9	21	AAK39702
Human cDNA for nov	520	12.9	23	AAK83457
Human breast and o	561	12.9	24	ABK34531
Human polynucleoti	814	12.9	21	AAF21759
C glutamicum codin	817	12.9	24	ABL90239
Arabidopsis thalia	1077	12.9	22	AAH66004
Arabidopsis thalia	1121	12.9	21	AAC50454
Arabidopsis thalia	1123	12.9	21	AAK38452
Arabidopsis thalia	1179	12.9	21	AAC50445
Corynebacterium gl	1185	12.9	21	AAC37125
DNA encoding novel	1200	12.9	22	AAF67975
Drosophila melanog	1355	12.9	23	AAK78569
DNA encoding novel	1616	12.9	23	ABL23849
Human polynucleoti	1657	12.9	23	AAK78567
DNA encoding novel	1704	12.9	24	ABL91706
Human transforming	1828	12.9	23	AAK87920
TGF-beta1 receptor	1918	12.9	17	AAK27593
Human Transforming	2090	12.9	14	AAQ42383
	2090	12.9	21	AAZ48864

83	17	12.9	2090	22	AAH27095	Human transforming	156	12.1	1232	24	ABN86297	Human protease cDN
84	17	12.9	2095	24	ABN86290	Type II TGFbeta re	157	12.1	1239	22	AAK51749	Human polynucleoti
85	17	12.9	2090	15	AAQ64830	TGF-beta receptor	c 158	12.1	1257	23	AAK51581	DNA encoding novel
86	17	12.9	2715	18	ABN79809	Fungal ZRC gene se	159	12.1	1312	21	AAK59439	Nucleotide sequenc
87	17	12.9	4590	18	AAAT94612	Rat penile neurona	160	12.1	1320	23	AAK58533	DNA encoding novel
88	17	12.9	5108	14	AAQ53403	Sequence encoding	161	12.1	1320	23	AAK58221	DNA encoding novel
89	17	12.9	5759	24	ABQ88231	Human osteoblast d	162	12.1	1385	22	AAH16077	Human cDNA sequenc
90	17	12.9	5759	24	ABK83678	Human cDNA differe	163	12.1	1406	16	AAQ97295	Marek's disease vi
91	17	12.9	5759	24	ABL66691	Lung cancer relate	164	12.1	1416	22	AAI60976	Human polynucleoti
92	17	12.9	5760	23	AAK76568	DNA encoding novel	165	12.1	1453	22	AAK85078	Atherosclerosis-as
93	17	12.9	13026	22	ABA15222	Human nervous syst	166	12.1	1464	24	ABA93723	Human amygdala der
94	17	12.9	13026	22	AAAL07529	Human reproductive	167	12.1	1492	22	ABA43542	Human breast cell
95	17	12.9	14970	22	AAAL07529	Human reproductive	168	12.1	1492	22	ABA43542	Probe #2209 for ge
96	17	12.9	14976	22	AAAS42037	Genomic sequence #	169	12.1	1492	22	AAK02262	Human brain expres
97	17	12.9	34980	22	AAH42038	Genomic sequence #	170	12.1	1492	22	AAK27709	Human bone marrow
98	16	12.1	60	24	AAH68527	C glutamic codin	171	12.1	1492	22	AAI12288	Probe #2231 for ge
99	16	12.1	134	21	ABN36497	Human spliced tran	172	12.1	1492	22	AAI13645	Probe #2331 used t
100	16	12.1	281	22	AAK98664	Human colon cancer	173	12.1	1492	22	AAI02204	Probe #2195 used t
101	16	12.1	294	22	AAAL25409	Human breast cance	174	12.1	1492	22	AAI02204	Human genome-deriv
102	16	12.1	307	22	AAAL16566	Human breast cance	175	12.1	1584	22	ABD02170	Human novel protei
103	16	12.1	311	22	AAH72404	Human cervical can	176	12.1	1584	22	ABD16566	DNA #58 encoding h
104	16	12.1	329	22	AAH70853	Human cervical can	177	12.1	1931	24	AAK30581	Human protease, PR
105	16	12.1	331	22	AAAL07663	Human breast cance	c 178	12.1	2109	22	AAK94222	Human full-length
106	16	12.1	331	22	AAH69518	Human cervical can	179	12.1	2135	23	AAK78501	DNA encoding novel
107	16	12.1	347	22	AAAL23252	Human breast cance	c 180	12.1	2199	19	AAK15431	Coding region of D
108	16	12.1	397	22	AAAL16301	Human breast cance	181	12.1	2199	19	AAK15431	Human secreted pro
109	16	12.1	426	22	AAH72979	Human cervical can	182	12.1	2352	21	AAK59695	cDNA from clone b1
110	16	12.1	429	24	ABL62160	Colon adenocarcino	183	12.1	2355	19	AAK63188	Human polynucleoti
111	16	12.1	434	23	ABK43038	Genomic sequence #	184	12.1	2355	24	ABQ92038	Human cDNA sequenc
112	16	12.1	444	24	ABN93633	Gene #131 used to	185	12.1	2527	22	AAH16264	Human cDNA sequenc
113	16	12.1	444	24	ABL64166	Stomach cancer rel	186	12.1	2545	22	AAH17775	Human novel protei
114	16	12.1	460	23	ABK42068	CDNA encoding nove	187	12.1	2693	22	AAI16565	Human novel protei
115	16	12.1	475	22	ABA51709	Human foetal liver	188	12.1	2693	22	AAI16568	Human novel protei
116	16	12.1	475	22	ABA51806	Human foetal liver	c 189	12.1	2693	22	AAK35107	DNA #57 encoding h
117	16	12.1	475	22	ABA21546	Probe #12 for gene	c 190	12.1	2693	22	AAK35109	DNA #59 encoding h
118	16	12.1	475	22	ABA21635	Probe #101 for gen	c 191	12.1	2693	22	AAK35110	DNA #60 encoding h
119	16	12.1	475	22	AAK00021	Human brain expres	c 192	12.1	2805	23	ABL05310	Drosophila melanog
120	16	12.1	475	22	AAK00113	Human bone marrow	c 193	12.1	2817	19	AAK15430	DNA encoding treha
121	16	12.1	475	22	AAK25456	Human bone marrow	194	12.1	2863	22	AAH48299	Human phosphoenol
122	16	12.1	475	22	AAK25550	Probe #12 for gene	195	12.1	3149	24	ABD39131	Rat obesity-specif
123	16	12.1	475	22	AAI10079	Probe #106 for gen	196	12.1	3270	19	AAV73472	Human UBP cDNA #2
124	16	12.1	475	22	AAI10173	Probe #13 used to	197	12.1	3270	23	AAK87519	DNA encoding novel
125	16	12.1	475	22	AAI31327	Probe #109 used to	c 198	12.1	3327	23	ABL11284	Drosophila melanog
126	16	12.1	475	22	AAI31423	Probe #13 used to	199	12.1	3437	16	AAK03437	Human matrix metal
127	16	12.1	475	22	AAI00022	Probe #107 used to	200	12.1	3437	24	ABL65860	Lung cancer relate
128	16	12.1	475	24	AAI000116	Human genome-deriv	201	12.1	3437	24	ABL35514	Human endometrial
129	16	12.1	475	24	ABK00021	Human genome-deriv	202	12.1	3456	16	AAK03436	Mouse laminin-15 g
130	16	12.1	475	24	ABK00119	Human breast cance	203	12.1	3540	24	AAI70820	Human breast cance
131	16	12.1	527	22	AAAL23305	Human cancer relat	204	12.1	4010	22	AAI26509	CDNA encoding huma
132	16	12.1	534	22	ABN60577	Human breast cance	205	12.1	4023	24	ABK14799	Drosophila melanog
133	16	12.1	567	22	AAAL14443	Human breast cance	c 206	12.1	4173	23	ABL05334	Pasteurella multoc
134	16	12.1	649	24	ABQ59067	Human colon cancer	c 207	12.1	4380	10	AAK91839	Pasteurella multoc
135	16	12.1	665	22	AAAL16538	Human novel protei	c 208	12.1	4381	12	AAQ10519	DNA encoding novel
136	16	12.1	665	22	AAK34878	CDNA encoding nove	209	12.1	4672	23	AAK37772	DNA encoding novel
137	16	12.1	685	22	AAH05932	Human cDNA clone (210	12.1	4815	24	ABK31757	Pseudomonas aerugi
138	16	12.1	712	21	AAK34195	Arabidopsis thalia	211	12.1	4863	23	AAK51489	CDNA encoding huma
139	16	12.1	733	22	AAH06563	Human cDNA clone (212	12.1	4890	20	AAK59766	Human cDNA differe
140	16	12.1	741	23	ABL05311	Drosophila melanog	213	12.1	4946	24	ABK3922	Human ORF ORF681
141	16	12.1	779	20	AAZ16506	Human gene expres	c 214	12.1	4998	22	AAK65573	Human ORF ORF681
142	16	12.1	802	24	ABQ89923	Human prostate exp	215	12.1	5300	21	AAK75126	Human reproductive
143	16	12.1	870	23	AAK53013	Enterococcus faeca	216	12.1	6093	22	AAAL05402	Human testicular a
144	16	12.1	977	24	ABQ73799	Human colon specif	217	12.1	6093	23	ABL98264	Drosophila melanog
145	16	12.1	1020	24	ABL90381	Human polynucleoti	218	12.1	6273	23	ABL15731	Drosophila melanog
146	16	12.1	1093	22	ABA48654	Human breast cell	c 219	12.1	6315	22	AAAL5744	Quorum sensing con
147	16	12.1	1093	22	ABA33630	Probe #12096 for g	c 220	12.1	6390	22	AAH81370	Genomic sequence #
148	16	12.1	1093	22	AAK14990	Human brain expres	c 221	12.1	6640	23	ABK42190	Gene #1724 used to
149	16	12.1	1093	22	AAK40723	Human bone marrow	222	12.1	6940	24	ABN95226	DNA encoding novel
150	16	12.1	1093	22	AAI21485	Probe #11418 for g	223	12.1	7856	23	AAK85531	DNA encoding novel
151	16	12.1	1093	22	AAI46777	Probe #15463 used	224	12.1	7856	23	AAK88213	DNA encoding novel
152	16	12.1	1093	22	AAI07183	Probe #7174 used t	225	12.1	7878	24	ABL84684	Stomach cancer rel
153	16	12.1	1093	22	ABSL14683	Human genome-deriv	c 226	12.1	8134	23	AAK03228	Drosophila melanog
154	16	12.1	1179	22	AAK52733	Human polynucleoti	227	12.1	8180	23	AAK88212	DNA encoding novel
155	16	12.1	1211	21	AAA62055	Hydrophobic domain	228	12.1	8201	17	AAI12170	pJG4-5-CDK-BP CDNA

229	16	12.1	8284	23	AA585535	DNA encoding novel	15	11.4	461	24	ABS07273	Human genome-deriv
230	16	12.1	8308	22	AA159190	Human polynucleoti	15	11.4	463	22	AAAL19114	Human breast cance
231	16	12.1	8409	22	AA157848	Human polynucleoti	15	11.4	471	21	AAAC00538	Human secreted pro
232	16	12.1	8420	22	AA157847	Human polynucleoti	15	11.4	472	23	ABV45615	Human prostate exp
233	16	12.1	8435	22	AA159633	Human polynucleoti	15	11.4	480	22	AAH06068	Human cDNA clone (
234	16	12.1	8435	22	AA159634	Human polynucleoti	15	11.4	485	22	AAH08949	Human cDNA clone (
235	16	12.1	8974	23	ABK42191	Genomic sequence #	15	11.4	487	23	AA579710	DNA encoding novel
236	16	12.1	10855	23	ABL15730	Drosophila melanog	15	11.4	499	21	AAAC40972	zea mays DNA fragm
237	16	12.1	10923	23	ABK42192	Genomic sequence #	15	11.4	505	22	AAH29674	Drosophila melanog
238	16	12.1	14822	20	AAK20543	Polynucleotide seq	15	11.4	508	22	AAAL13213	Human breast cance
239	16	12.1	16956	21	AA291923	Wild type (C57BL/6	15	11.4	511	24	ABL83055	Human ovarian cenc
240	16	12.1	22680	22	AAK66308	Human immune/haema	15	11.4	518	22	AA531184	Human diagnostic a
241	16	12.1	22680	22	AAK73334	Human immune/haema	15	11.4	520	22	AAAL22083	Human breast cance
242	16	12.1	22680	22	AAK73344	Human immune/haema	15	11.4	522	24	ABN66905	Streptococcus poly
243	16	12.1	22680	22	AAK73625	Human immune/haema	15	11.4	526	24	ABK39435	DNA encoding lung
244	16	12.1	22680	22	AAK73847	Human immune/haema	15	11.4	533	23	AAH13579	Human cDNA clone (
245	16	12.1	22680	22	AAK73934	Human immune/haema	15	11.4	536	24	ABK09684	Human ovarian tumo
246	16	12.1	22680	22	AAK78350	Human immune/haema	15	11.4	547	22	AAH05727	Human cDNA Clone (
247	16	12.1	37716	23	AA595953	Propionibacterium	15	11.4	554	22	AAAL21812	Human breast cance
248	16	12.1	46765	24	AA599306	DNA encoding Aldeh	15	11.4	555	22	AAAL21487	Human breast cance
249	16	12.1	112190	22	AAH44801	Human GPCR protein	15	11.4	557	22	AAAL12614	Human breast cance
250	16	12.1	368004	24	ABL57909	Human transporter	15	11.4	563	23	ABV06749	Human prostate exp
251	15	11.4	27	22	AA504265	Saccharomyces cere	15	11.4	566	23	ABV54753	Human prostate exp
252	15	11.4	27	24	AAK53065	S. Cerevisiae HMG	15	11.4	574	24	ABN65854	Human cancer relat
253	15	11.4	45	22	ADL17288	Human thymosin bet	15	11.4	580	23	ABV06926	Human prostate exp
254	15	11.4	213	16	AAAT01156	Human gene signatu	15	11.4	581	22	AAH09232	Human cDNA clone (
255	15	11.4	231	22	AA560131	Human cancer agent	15	11.4	584	22	AAH05114	Human cDNA clone (
256	15	11.4	257	21	AAAM33117	Human secreted exp	15	11.4	595	22	AAK69867	Human immune/haema
257	15	11.4	267	21	AAAC00908	Human secreted pro	15	11.4	597	24	ABN65484	Human cancer relat
258	15	11.4	273	16	AAAT24694	Human gene signatu	15	11.4	615	23	ABV54822	Human prostate exp
259	15	11.4	277	20	AAAL13372	Human gene express	15	11.4	617	21	AAAF09329	Fusarium venenatum
260	15	11.4	293	22	AA560106	Human cancer agent	15	11.4	647	22	AAAL12941	Human breast cance
261	15	11.4	293	22	AAH77501	Human DNA repair p	15	11.4	653	23	ABV36695	Human prostate exp
262	15	11.4	304	22	AAI128952	Colon tumour relat	15	11.4	654	23	ABV36870	Human prostate exp
263	15	11.4	338	21	AAAC53477	Arabidopsis thalia	15	11.4	670	23	AA576415	DNA encoding novel
264	15	11.4	347	22	AAAL05011	Human reproductive	15	11.4	674	24	AA561600	Lung small cell ca
265	15	11.4	347	23	ABL57904	Human testicular a	15	11.4	681	22	AAAL17184	Human zinculine fr
266	15	11.4	349	22	AAAL01333	Human reproductive	15	11.4	684	22	AAH05689	Human cDNA clone (
267	15	11.4	349	23	ABL96787	Human testicular a	15	11.4	696	21	AAH05469	E. coli essential
268	15	11.4	351	22	ABA066677	Human cDNA SEQ ID	15	11.4	696	21	AAH8696	E. coli FUN essent
269	15	11.4	366	22	AA560497	Human cancer agent	15	11.4	696	22	AAH81404	Escherichia coli p
270	15	11.4	366	24	ABQ58932	Human colon cancer	15	11.4	696	23	AA552487	E. coli DNA for ce
271	15	11.4	369	14	AAO61270	Human brain Expres	15	11.4	714	23	AA587333	DNA encoding novel
272	15	11.4	374	22	AAAL00541	Human reproductive	15	11.4	714	23	AA592876	DNA encoding novel
273	15	11.4	388	24	ABN600045	Human cancer relat	15	11.4	715	20	AAH22212	Human secreted pro
274	15	11.4	390	22	AAI92498	Human polynucleoti	15	11.4	724	22	AAH05993	Human cDNA clone (
275	15	11.4	396	22	AAAF94957	Human ovarian cenc	15	11.4	724	24	AA562054	Porcine muscular s
276	15	11.4	396	24	ABT03224	Human ovarian carc	15	11.4	781	22	AAI96499	Human neuroblastom
277	15	11.4	396	24	ABL48907	Ovarian carcinoma	15	11.4	782	24	ABN96994	Gene #3492 used to
278	15	11.4	399	21	AAH30571	Human colon cancer	15	11.4	786	20	AAZ16704	Human gene express
279	15	11.4	399	22	AAH12265	Human cDNA clone (15	11.4	817	22	AAH04312	Human cDNA clone (
280	15	11.4	406	23	ABV14058	Human prostate exp	15	11.4	837	22	AAF80410	Nucleotide sequenc
281	15	11.4	419	23	AA592468	DNA encoding novel	15	11.4	837	24	ABK65175	Arabidopsis cDNA e
282	15	11.4	422	14	AAQ59867	Human brain Expres	15	11.4	843	21	AA537750	Arabidopsis thalia
283	15	11.4	425	22	AAH12338	Human cDNA clone (15	11.4	849	23	AA588546	DNA encoding novel
284	15	11.4	441	23	ABV04889	Human prostate exp	15	11.4	867	24	ABN67373	Streptococcus poly
285	15	11.4	442	22	AAAF97958	Human secreted pro	15	11.4	888	23	ABL09165	Drosophila melanog
286	15	11.4	443	23	ABV35163	Human prostate exp	15	11.4	889	18	AAAT75097	Arabidopsis flower
287	15	11.4	443	23	ABV44005	Human prostate exp	15	11.4	908	22	AAAL02651	Human reproductive
288	15	11.4	446	24	ABK55170	Human colon cancer	15	11.4	908	23	ABL97318	Human testicular a
289	15	11.4	448	22	AAH33334	Human colon cancer	15	11.4	913	20	AAZ28273	Rat neuronal lmed
290	15	11.4	459	23	ABV15813	Human prostate exp	15	11.4	913	22	AAH05068	Human cDNA clone (
291	15	11.4	460	20	AAZ24593	Human lung tumor a	15	11.4	918	23	ABL03121	Drosophila melanog
292	15	11.4	460	21	AAAC65822	Human lung cancer-	15	11.4	930	22	AA530994	Human diagnostic a
293	15	11.4	460	24	ABL49041	Human lung tumour	15	11.4	936	23	ABV22483	Human prostate exp
294	15	11.4	461	22	ABA58579	Human foetal liver	15	11.4	936	23	ABV23104	Human prostate exp
295	15	11.4	461	22	ABA27600	Probe #6066 for ge	15	11.4	936	23	ABV28298	Human prostate exp
296	15	11.4	461	22	AAK06701	Human brain expres	15	11.4	936	23	ABV28942	Human prostate exp
297	15	11.4	461	22	AAK32400	Human bone marrow	15	11.4	975	18	AAT68791	Melanocortin-5 rec
298	15	11.4	461	22	AAK32488	Human bone marrow	15	11.4	975	19	AAV62353	Melanocortin-5 rec
299	15	11.4	461	22	AAI16060	Probe #5993 for ge	15	11.4	975	19	AAV03916	Mouse melanocortin
300	15	11.4	461	22	AAI38243	Probe #6929 used t	15	11.4	975	19	AAV06405	Mouse melanocortin
301	15	11.4	461	24	ABS07184	Human genome-deriv	15	11.4	975	20	AAH01965	Mouse MC5 DNA. Mu

c 375	15	11.4	975	21	AAV15830	DNA encoding a hum	448	15	11.4	2490	22	AAH17721	Human cDNA sequenc
c 376	15	11.4	978	19	AAV19143	Mouse melanocortin	449	15	11.4	2506	22	AAH44993	Human encoding nove
c 377	15	11.4	978	20	AAV63708	Mouse melanocortin	450	15	11.4	2509	16	AAO92776	Human thymopoietin
c 378	15	11.4	979	23	ABV22327	Human prostate exp	451	15	11.4	2512	22	AAH17639	Human cDNA sequenc
c 379	15	11.4	979	23	ABV22327	Human prostate exp	452	15	11.4	2512	22	AAH17639	Human polynucleoti
c 380	15	11.4	1014	22	AAH44571	Human prostate exp	453	15	11.4	2571	21	AAO16123	Human polynucleoti
c 381	15	11.4	1014	22	AAH44571	Human prostate exp	454	15	11.4	2571	21	AAO16123	Human pancreatic c
c 382	15	11.4	1032	24	AAH98422	Human cDNA encodin	455	15	11.4	2615	22	AAO2597	Human Electraon tra
c 383	15	11.4	1032	24	AAH98422	Human cDNA encodin	456	15	11.4	2621	22	AAH15693	Human cDNA sequenc
c 384	15	11.4	1093	21	AAH56537	Arabidopsis thalia	457	15	11.4	2638	24	ABN59719	Novel human coding
c 385	15	11.4	1093	21	AAH56537	Arabidopsis thalia	458	15	11.4	2638	24	ABN59719	DNA encoding novel
c 386	15	11.4	1128	22	AAH56537	Human cytokine fam	459	15	11.4	2676	23	AAH56537	Human hydrophobic
c 387	15	11.4	1132	22	AAH56537	DNA encoding nove	460	15	11.4	2694	22	AAH56537	Human cDNA encodin
c 388	15	11.4	1161	21	AAH56537	cDNA encoding nove	461	15	11.4	2760	22	AAH56537	Human prostate exp
c 389	15	11.4	1378	22	AAH15719	Human secreted pro	462	15	11.4	2826	23	ABV24329	Human prostate exp
c 390	15	11.4	1430	22	AAH14084	Human cDNA sequenc	463	15	11.4	2826	23	ABV24329	Drosophila melanog
c 391	15	11.4	1431	22	AAH14084	Human FCTR3a DNA s	464	15	11.4	2877	23	ABL02791	DNA encoding novel
c 392	15	11.4	1434	22	AAH14084	Human polynucleoti	465	15	11.4	2928	23	AAH16666	DNA encoding novel
c 393	15	11.4	1477	24	AAH14084	Human polynucleoti	466	15	11.4	3005	23	AAH16666	DNA encoding novel
c 394	15	11.4	1482	23	AAH14084	Human polynucleoti	467	15	11.4	3011	23	AAH16666	Human cDNA encodin
c 395	15	11.4	1499	22	AAH14084	Human polynucleoti	468	15	11.4	3012	23	AAH16666	Human cDNA encodin
c 396	15	11.4	1526	22	AAH14084	Human polynucleoti	469	15	11.4	3027	24	AAH16666	Human cDNA encodin
c 397	15	11.4	1562	22	AAH14084	Human polynucleoti	470	15	11.4	3102	24	AAH16666	Human cDNA encodin
c 398	15	11.4	1586	19	AAH14084	Human polynucleoti	471	15	11.4	3116	22	AAH16666	Human cDNA encodin
c 399	15	11.4	1587	22	AAH14084	Human polynucleoti	472	15	11.4	3126	20	AAH16666	Human cDNA encodin
c 400	15	11.4	1592	22	AAH14084	Human polynucleoti	473	15	11.4	3126	20	AAH16666	Human cDNA encodin
c 401	15	11.4	1622	24	AAH14084	Human polynucleoti	474	15	11.4	3201	22	AAH16666	Human cDNA encodin
c 402	15	11.4	1626	20	AAH14084	Human polynucleoti	475	15	11.4	3201	22	AAH16666	Human cDNA encodin
c 403	15	11.4	1656	22	AAH14084	Human polynucleoti	476	15	11.4	3276	22	AAH16666	Human cDNA encodin
c 404	15	11.4	1674	22	AAH14084	Human polynucleoti	477	15	11.4	3302	23	ABL09164	Human cDNA encodin
c 405	15	11.4	1674	22	AAH14084	Human polynucleoti	478	15	11.4	3367	23	ABL24464	Drosophila melanog
c 406	15	11.4	1674	22	AAH14084	Human polynucleoti	479	15	11.4	3465	22	ABH18323	Human nervous syst
c 407	15	11.4	1674	22	AAH14084	Human polynucleoti	480	15	11.4	3507	22	AAH18323	Human nervous syst
c 408	15	11.4	1727	22	AAH14084	Human polynucleoti	481	15	11.4	3507	22	AAH18323	Human nervous syst
c 409	15	11.4	1727	22	AAH14084	Human polynucleoti	482	15	11.4	3510	22	AAH18323	Human nervous syst
c 410	15	11.4	1728	20	AAH14084	Human polynucleoti	483	15	11.4	3510	22	AAH18323	Human nervous syst
c 411	15	11.4	1743	22	AAH14084	Human polynucleoti	484	15	11.4	3510	22	AAH18323	Human nervous syst
c 412	15	11.4	1754	24	AAH14084	Human polynucleoti	485	15	11.4	3510	22	AAH18323	Human nervous syst
c 413	15	11.4	1853	22	AAH14084	Human polynucleoti	486	15	11.4	3510	22	AAH18323	Human nervous syst
c 414	15	11.4	1854	22	AAH14084	Human polynucleoti	487	15	11.4	3510	22	AAH18323	Human nervous syst
c 415	15	11.4	1854	22	AAH14084	Human polynucleoti	488	15	11.4	3510	22	AAH18323	Human nervous syst
c 416	15	11.4	1893	22	AAH14084	Human polynucleoti	489	15	11.4	3510	22	AAH18323	Human nervous syst
c 417	15	11.4	1897	19	AAH14084	Human polynucleoti	490	15	11.4	3510	22	AAH18323	Human nervous syst
c 418	15	11.4	1916	23	AAH14084	Human polynucleoti	491	15	11.4	3510	22	AAH18323	Human nervous syst
c 419	15	11.4	1964	22	AAH14084	Human polynucleoti	492	15	11.4	3510	22	AAH18323	Human nervous syst
c 420	15	11.4	1984	21	AAH14084	Human polynucleoti	493	15	11.4	3510	22	AAH18323	Human nervous syst
c 421	15	11.4	1985	22	AAH14084	Human polynucleoti	494	15	11.4	3510	22	AAH18323	Human nervous syst
c 422	15	11.4	2012	19	AAH14084	Human polynucleoti	495	15	11.4	3510	22	AAH18323	Human nervous syst
c 423	15	11.4	2012	19	AAH14084	Human polynucleoti	496	15	11.4	3510	22	AAH18323	Human nervous syst
c 424	15	11.4	2031	22	AAH14084	Human polynucleoti	497	15	11.4	3510	22	AAH18323	Human nervous syst
c 425	15	11.4	2040	22	AAH14084	Human polynucleoti	498	15	11.4	3510	22	AAH18323	Human nervous syst
c 426	15	11.4	2051	22	AAH14084	Human polynucleoti	499	15	11.4	3510	22	AAH18323	Human nervous syst
c 427	15	11.4	2051	22	AAH14084	Human polynucleoti	500	15	11.4	3510	22	AAH18323	Human nervous syst
c 428	15	11.4	2051	22	AAH14084	Human polynucleoti	501	15	11.4	3510	22	AAH18323	Human nervous syst
c 429	15	11.4	2053	21	AAH14084	Human polynucleoti	502	15	11.4	3510	22	AAH18323	Human nervous syst
c 430	15	11.4	2070	24	AAH14084	Human polynucleoti	503	15	11.4	3510	22	AAH18323	Human nervous syst
c 431	15	11.4	2088	21	AAH14084	Human polynucleoti	504	15	11.4	3510	22	AAH18323	Human nervous syst
c 432	15	11.4	2097	22	AAH14084	Human polynucleoti	505	15	11.4	3510	22	AAH18323	Human nervous syst
c 433	15	11.4	2097	22	AAH14084	Human polynucleoti	506	15	11.4	3510	22	AAH18323	Human nervous syst
c 434	15	11.4	2190	22	AAH14084	Human polynucleoti	507	15	11.4	3510	22	AAH18323	Human nervous syst
c 435	15	11.4	2195	22	AAH14084	Human polynucleoti	508	15	11.4	3510	22	AAH18323	Human nervous syst
c 436	15	11.4	2206	16	AAH14084	Human polynucleoti	509	15	11.4	3510	22	AAH18323	Human nervous syst
c 437	15	11.4	2208	22	AAH14084	Human polynucleoti	510	15	11.4	3510	22	AAH18323	Human nervous syst
c 438	15	11.4	2215	23	AAH14084	Human polynucleoti	511	15	11.4	3510	22	AAH18323	Human nervous syst
c 439	15	11.4	2219	23	AAH14084	Human polynucleoti	512	15	11.4	3510	22	AAH18323	Human nervous syst
c 440	15	11.4	2220	23	AAH14084	Human polynucleoti	513	15	11.4	3510	22	AAH18323	Human nervous syst
c 441	15	11.4	2235	23	AAH14084	Human polynucleoti	514	15	11.4	3510	22	AAH18323	Human nervous syst
c 442	15	11.4	2338	22	AAH14084	Human polynucleoti	515	15	11.4	3510	22	AAH18323	Human nervous syst
c 443	15	11.4	2344	24	AAH14084	Human polynucleoti	516	15	11.4	3510	22	AAH18323	Human nervous syst
c 444	15	11.4	2345	22	AAH14084	Human polynucleoti	517	15	11.4	3510	22	AAH18323	Human nervous syst
c 445	15	11.4	2403	22	AAH14084	Human polynucleoti	518	15	11.4	3510	22	AAH18323	Human nervous syst
c 446	15	11.4	2436	23	AAH14084	Human polynucleoti	519	15	11.4	3510	22	AAH18323	Human nervous syst
c 447	15	11.4	2436	23	AAH14084	Human polynucleoti	520	15	11.4	3510	22	AAH18323	Human nervous syst

c 521	15	11.4	7120	23	ABL04480	Drosophila melanog	c 594	14	10.6	15	AAQ69677	Human hepatic lipa
c 522	15	11.4	7676	19	AAV62906	Human galactokinas	c 595	14	10.6	51	AAT64139	Human hepatic lipa
c 523	15	11.4	7727	22	AAV04997	Human reproductiv	c 596	14	10.6	51	AAK17427	Test sequence from
c 524	15	11.4	7727	23	ABL97890	Human testicular a	c 597	14	10.6	51	ABK62918	DNA binding molecu
c 525	15	11.4	7997	23	ABL28198	Drosophila melanog	c 598	14	10.6	60	ABN32701	Human spliced tran
c 526	15	11.4	8095	24	ABL95730	Gene #2228 used to	c 599	14	10.6	60	ABN33731	Human spliced tran
c 527	15	11.4	8280	23	AAW74108	DNA encoding novel	c 600	14	10.6	60	ABN45615	Human spliced tran
c 528	15	11.4	8367	18	AAW73721	Human filamin cDNA	c 601	14	10.6	65	ABN53425	Mouse spliced tran
c 529	15	11.4	8367	20	AAZ06508	Carboxyl terminal	c 602	14	10.6	81	ABA36518	Probe #14984 for g
c 530	15	11.4	8368	22	AAH81766	Human differential	c 603	14	10.6	99	AAW48762	Polyclonal anti-fe
c 531	15	11.4	8368	24	ABL62478	Colon adenocarcino	c 604	14	10.6	122	ABA72504	Human foetal liver
c 532	15	11.4	8806	17	AAT09349	Rat IP3 receptor g	c 605	14	10.6	122	ABA38258	Probe #16724 for g
c 533	15	11.4	9729	22	AAAI4089	Human FCTR3f DNA s	c 606	14	10.6	122	AAK30927	Human brain expres
c 534	15	11.4	9826	22	AAAI4085	Human FCTR3b DNA s	c 607	14	10.6	122	AAK47080	Human bone marrow
c 535	15	11.4	10319	22	AAAS27632	DNA encoding novel	c 608	14	10.6	122	AAI52916	Probe #21602 used
c 536	15	11.4	10438	23	AAW74106	DNA encoding novel	c 609	14	10.6	122	ABS21314	Human genome-deriv
c 537	15	11.4	10479	23	ABL15220	Drosophila melanog	c 610	14	10.6	137	ABA50761	Human breast cell
c 538	15	11.4	10718	14	AAQ51476	DEN1-S275/90 (ECAC	c 611	14	10.6	137	ABA68729	Human foetal liver
c 539	15	11.4	12081	23	ABL09560	Drosophila melanog	c 612	14	10.6	137	ABA35692	Probe #14158 for g
c 540	15	11.4	12666	19	AAV52270	Streptococcus pneu	c 613	14	10.6	137	AAK17072	Human brain expres
c 541	15	11.4	15571	24	ABN95889	Gene #2387 used to	c 614	14	10.6	137	AAK42855	Human bone marrow
c 542	15	11.4	15571	24	ABL68256	Kidney cancer rela	c 615	14	10.6	137	AAI23619	Probe #13552 for g
c 543	15	11.4	15571	24	ABL68528	Kidney cancer rela	c 616	14	10.6	137	AAI48930	Probe #17616 used
c 544	15	11.4	15848	20	AAZ32190	Human heparin cofa	c 617	14	10.6	137	AAI09234	Probe #9225 used t
c 545	15	11.4	15849	24	ABN93864	Gene #2362 used to	c 618	14	10.6	137	ABSI6909	Human genome-deriv
c 546	15	11.4	16428	22	AAK81010	Human immune/haema	c 619	14	10.6	138	ABV04928	Human prostate exp
c 547	15	11.4	16428	23	ABK42726	Genomic sequence #	c 620	14	10.6	140	AAT20692	Human gene signatu
c 548	15	11.4	18909	23	ABL07204	Drosophila melanog	c 621	14	10.6	140	AAI00065	Human reproductiv
c 549	15	11.4	19736	24	ABN89474	Human protease gen	c 622	14	10.6	151	AAW2978	Human foetal liver
c 550	15	11.4	19736	24	ABA92158	Human protease gen	c 623	14	10.6	151	AAK21410	Human brain expres
c 551	15	11.4	20029	23	ABK42737	Genomic sequence #	c 624	14	10.6	151	AAK47570	Human bone marrow
c 552	15	11.4	23241	22	AAK84291	Human immune/haema	c 625	14	10.6	151	AAI53402	Probe #22088 used
c 553	15	11.4	23241	22	AAK87225	Human immune/haema	c 626	14	10.6	157	ABA74875	Human foetal liver
c 554	15	11.4	28720	23	ABL07302	Drosophila melanog	c 627	14	10.6	157	ABA39572	Probe #18038 for g
c 555	15	11.4	31834	22	AAK81009	Human immune/haema	c 628	14	10.6	157	AAK23368	Human brain expres
c 556	15	11.4	31834	22	ABK42725	Genomic sequence #	c 629	14	10.6	157	AAK49523	Human bone marrow
c 557	15	11.4	32035	22	AAK81008	Human immune/haema	c 630	14	10.6	157	AAI26635	Probe #16568 for g
c 558	15	11.4	32035	22	ABK42724	Genomic sequence #	c 631	14	10.6	157	AAI55400	Probe #24086 used
c 559	15	11.4	32220	22	ABA20618	Human nervous syst	c 632	14	10.6	157	AAI23007	Human genome-deriv
c 560	15	11.4	41488	22	AAK87512	Human immune/haema	c 633	14	10.6	166	AAC29995	Human secreted pro
c 561	15	11.4	49634	24	ABL68647	Kidney cancer rela	c 634	14	10.6	175	ABL78254	Human ovarian canc
c 562	15	11.4	50575	23	ABL29244	Drosophila melanog	c 635	14	10.6	181	AAK72231	Human immune/haema
c 563	15	11.4	66788	23	ABQ88140	Human osteoblast d	c 636	14	10.6	191	ABA76564	Human foetal liver
c 564	15	11.4	66788	24	AAQ85915	Propionibacterium	c 637	14	10.6	191	ABA41076	Probe #19542 for g
c 565	15	11.4	76574	23	ABL03438	Drosophila melanog	c 638	14	10.6	191	AAK52505	Human brain expres
c 566	15	11.4	94618	21	AAW22285	BAC containing rep	c 639	14	10.6	191	AAK51208	Human bone marrow
c 567	15	11.4	95833	21	AAW22297	BAC containing rep	c 640	14	10.6	191	AAI28220	Probe #18153 for g
c 568	15	11.4	99960	21	AAZ50905	Human TBC-1 partia	c 641	14	10.6	191	AAI57262	Probe #25948 used
c 569	15	11.4	101786	21	AAW22293	BAC containing rep	c 642	14	10.6	191	ABS24743	Human genome-deriv
c 570	15	11.4	160755	23	AAH88704	Human DNA sequence	c 643	14	10.6	193	ABA51494	Human breast cell
c 571	15	11.4	160771	24	ABQ88179	Human osteoblast d	c 644	14	10.6	193	ABA36450	Probe #14916 for g
c 572	15	11.4	197496	24	ABN85584	Human EGFR SEQ ID	c 645	14	10.6	193	AAK17771	Human brain expres
c 573	15	11.4	215980	24	AAI38337	Complementary stra	c 646	14	10.6	193	AAK43599	Human bone marrow
c 574	15	11.4	305107	22	AAH62689	Shrimp white spot	c 647	14	10.6	193	AAI24382	Probe #14315 for g
c 575	15	11.4	335913	22	AAI61371	Soybean 240017 reg	c 648	14	10.6	193	AAI09920	Probe #9911 used t
c 576	15	11.4	335913	22	AAI61372	Soybean 240017 reg	c 649	14	10.6	193	ABS17764	Human genome-deriv
c 577	15	11.4	910715	20	AAZ20248	Borrelia burgdorfe	c 650	14	10.6	195	ABK97946	Botulinum toxin ge
c 578	15	11.4	1038602	20	AAZ01425	Complete genome se	c 651	14	10.6	198	AAI56087	Human foetal liver
c 579	15	11.4	2365589	24	ABA90521	Genomic sequence o	c 652	14	10.6	198	ABA40132	Probe #18598 for g
c 580	15	11.4	2365589	24	ABA90521	Genomic sequence o	c 653	14	10.6	198	AAK24085	Human brain expres
c 581	15	11.4	2944528	24	ABA03041	Listeria monocytog	c 654	14	10.6	198	AAK50147	Human bone marrow
c 582	15	11.4	4403765	22	AAI99683	Mycobacterium tube	c 655	14	10.6	198	AAI27230	Probe #17163 for g
c 583	15	11.4	4411529	22	AAI99682	Mycobacterium tube	c 656	14	10.6	198	AAI27230	Probe #24773 used
c 584	14	10.6	27	24	ABK97949	Botulin toxin gene	c 657	14	10.6	198	ABL27031	Drosophila melanog
c 585	14	10.6	27	24	ABA91170	Physcomitrella pat	c 658	14	10.6	198	ABS19180	Human genome-deriv
c 586	14	10.6	28	24	AAD28745	Human ion channel	c 659	14	10.6	198	ABS23624	Human genome-deriv
c 587	14	10.6	34	21	AAAS4350	Mutagenic primer E	c 660	14	10.6	198	AAD22386	Human OREX polynuc
c 588	14	10.6	38	23	AAK04417	Human NOGO Hammerh	c 661	14	10.6	201	AAD28727	Human ion channel
c 589	14	10.6	38	23	ABK07955	Human CD20 Hammerh	c 662	14	10.6	213	AAT13902	Human amygdaloid G
c 590	14	10.6	50	15	AAQ69676	Human hepatic lipa	c 663	14	10.6	213	AAW98798	3' fragment of G p
c 591	14	10.6	50	18	AAW64138	Human hepatic lipa	c 664	14	10.6	215	ABA73589	Human foetal liver
c 592	14	10.6	50	20	AAW17426	Test sequence from	c 665	14	10.6	215	ABA38846	Probe #17312 for g
c 593	14	10.6	50	24	ABK82917	DNA binding molecu	c 666	14	10.6	215	AAK22035	Human brain expres

c 667	14	10.6	215	22	AAK48198	Human bone marrow
c 668	14	10.6	215	22	AAI5A028	Probe #22714 used
c 669	14	10.6	215	24	ABS22089	Human genome-deriv
c 670	14	10.6	215	24	ABN25582	Human ORFX polynuc
c 671	14	10.6	224	16	AAT20613	Human gene signatu
c 672	14	10.6	226	21	AAC12797	Human secreted pro
c 673	14	10.6	233	21	AAZ58761	Human huntingtin-1
c 674	14	10.6	236	24	ABL83965	Human ovarian canc
c 675	14	10.6	249	19	AXL10948	Human biallelic po
c 676	14	10.6	256	21	AAZ22202	Single nucleotide
c 677	14	10.6	260	21	AAZ72193	Eucalyptus grandis
c 678	14	10.6	262	21	AAA67400	Drosophila melanog
c 679	14	10.6	262	22	AAH29389	Human ovarian tumo
c 680	14	10.6	263	20	AAZ77510	Human secreted pro
c 681	14	10.6	275	21	AAC30596	Human ORFX polynuc
c 682	14	10.6	279	24	ABN18171	Human brain Expres
c 683	14	10.6	281	14	AAQ60426	Human prostate exp
c 684	14	10.6	284	23	ABL73869	Corn tassal-deriv
c 685	14	10.6	290	24	ABL75554	Corn tassal-deriv
c 686	14	10.6	295	24	AAZ07474	Human secreted pro
c 687	14	10.6	298	21	AAC05655	Human secreted pro
c 688	14	10.6	301	21	AAC05655	Human brain Expres
c 689	14	10.6	304	14	AAQ59447	Human breast cell
c 690	14	10.6	311	22	ABA49017	Human prostate exp
c 691	14	10.6	311	22	ABA34018	Human breast cell
c 692	14	10.6	311	22	AAK15377	Human brain Expres
c 693	14	10.6	311	22	AAK11101	Human bone marrow
c 694	14	10.6	311	22	AAI21864	Probe #11797 for g
c 695	14	10.6	311	22	AAI47150	Probe #15836 used
c 696	14	10.6	311	22	AAI07552	Probe #7543 used t
c 697	14	10.6	311	23	ABV08977	Human genome-deriv
c 698	14	10.6	311	24	ABSL5093	Human prostate exp
c 699	14	10.6	312	23	ABSL5093	Drosophila melanog
c 700	14	10.6	318	23	ABV39131	Human prostate exp
c 701	14	10.6	321	22	AAH68301	C glutamicum codin
c 702	14	10.6	326	24	ABN21966	Human ORFX polynuc
c 703	14	10.6	329	22	ABAL4266	Human nervous syst
c 704	14	10.6	330	22	AAH35502	Human colon cancer
c 705	14	10.6	331	22	AAZ79981	Nucleotide sequenc
c 706	14	10.6	332	20	AAH87538	Human single nucle
c 707	14	10.6	333	15	ABV04932	Merlin gene fragme
c 708	14	10.6	343	24	ABL70733	Human prostate exp
c 709	14	10.6	343	24	ABV47297	Corn tassal-deriv
c 710	14	10.6	345	23	ABV47297	Human prostate exp
c 711	14	10.6	347	20	AAZ54700	Muscarinic acetyl
c 712	14	10.6	347	21	AAZ20269	Human muscarinic a
c 713	14	10.6	347	21	AAZ34147	Human adenosine re
c 714	14	10.6	347	22	AAZ0295	Human cancer agent
c 715	14	10.6	349	22	AAZ27606	Human -colon cancer
c 716	14	10.6	350	22	AAI80171	Human polynucleoti
c 717	14	10.6	352	23	ABV14101	Human prostate exp
c 718	14	10.6	353	16	AAT23525	Human gene signatu
c 719	14	10.6	356	24	ABN88217	Human colon cancer
c 720	14	10.6	357	14	AAQ61279	Human brain Expres
c 721	14	10.6	357	24	ABL81388	Human ovarian canc
c 722	14	10.6	359	24	ABQ58869	Human colon cancer
c 723	14	10.6	359	24	ABL37464	Human colon tumour
c 724	14	10.6	359	24	ABL37655	Human colon tumour
c 725	14	10.6	359	24	ABL37891	Human colon tumour
c 726	14	10.6	363	22	AAZ65606	Human colon cancer
c 727	14	10.6	363	23	ABV07621	Human prostate exp
c 728	14	10.6	369	22	AAI19455	Human bladder rela
c 729	14	10.6	369	22	AAI64100	CDNA encoding mri
c 730	14	10.6	372	21	AAZ61665	Skin cell CDNA, SE
c 731	14	10.6	372	22	AAZ99598	Murine CDNA isolat
c 732	14	10.6	372	24	ABL34750	Human excretory re
c 733	14	10.6	372	24	AAI198856	Human bladder rela
c 734	14	10.6	377	22	AAI64043	Novel human polynu
c 735	14	10.6	377	22	AAZ65134	Human secreted pro
c 736	14	10.6	381	21	AAC03970	Novel human diagno
c 737	14	10.6	381	22	AAZ53923	Human prostate exp
c 738	14	10.6	386	23	ABV57969	Human colon cancer
c 739	14	10.6	393	22	AAH35874	Human colon cancer

Nucleotide sequenc
Human prostate exp
Probe #5116 for ge
Human lung cancer
Novel human polynu
Staphylococcus aur
Human breast cell
Human foetal liver
Probe #3628 for ge
Human brain expres
Human bone marrow
Probe #3662 for ge
Probe #3773 used t
Probe #3398 used t
Gene #1098 used to
Human genome-deriv
Human cancer relat
Human neuroregulin g
Human neuroregulin g
Human neuroregulin g
Human immune/haema
Human prostate exp
Human prostate exp
Human secreted exp
Human ORFX polynuc
Human brain Expres
CDNA encoding nove
Human colon cancer
Human immune/haema
Colon adenocarcino
Human immune/haema
Human contig polyn
Human contig polyn
DNA encoding novel
Human polynucleoti
Human prostate exp
Colon adenocarcino
y lipolytica EST-d
DNA encoding novel
DNA encoding novel
DNA encoding novel
Human prostate exp
Human cervical can
Human breast cell
Human foetal liver
Probe #4259 for ge
Human brain expres
Human bone marrow
Probe #4350 for ge
Probe #4476 used t
Probe #4231 used t
Human genome-deriv
Human CDNA clone (
Human cancer relat
Human kidney relat
Human secreted pro
Human tumour assoc
Human prostate exp
Chromosome 6q27 va
Human polynucleoti
Human chromosome 6
Human nervous syst
Human breast cance
DNA encoding novel
Pinus radiata tran
Human foetal liver
Probe #619 for gen
Human brain expres
Human bone marrow

813	14	10.6	466	22	AA110708	Probe #641 for gen	886	14	10.6	533	22	AAH70263	Human cervical can
814	14	10.6	466	22	AAI31966	Probe #652 used to	887	14	10.6	533	24	ABN73668	Bovine embryonic g
815	14	10.6	466	22	AAI00638	Probe #629 used to	888	14	10.6	534	22	AAH13527	Human cDNA clone (
816	14	10.6	466	24	ABS00663	Human genome-deriv	c 889	14	10.6	535	21	AAA41276	Human secreted exp
c 817	14	10.6	471	22	ABA42624	Human foetal liver	c 890	14	10.6	535	24	ABN61678	Human cancer relat
c 818	14	10.6	471	22	ABA53053	Human foetal liver	c 891	14	10.6	536	23	ABV55416	Human prostate exp
c 819	14	10.6	471	22	ABA22827	Probe #1293 for ge	c 892	14	10.6	541	22	AAK36437	Human bone marrow
c 820	14	10.6	471	22	AAI32657	Probe #1343 used t	893	14	10.6	542	22	AAK86449	Human immune/haema
821	14	10.6	471	22	ABN69058	Streptococcus poly	894	14	10.6	542	23	ABK43053	Genomic sequence #
c 822	14	10.6	472	22	ABN43918	Human breast cell	895	14	10.6	543	23	ABK42073	CDNA encoding nove
c 823	14	10.6	472	22	ABA24150	Probe #2616 for ge	896	14	10.6	543	23	ABV60205	Human prostate exp
c 824	14	10.6	472	22	AAK02667	Human brain expres	c 897	14	10.6	544	23	AAK92492	Human cDNA 3'-end
c 825	14	10.6	472	22	AAK28104	Human bone marrow	898	14	10.6	545	22	AAK92496	CDNA encoding nove
c 826	14	10.6	472	22	AAI12680	Probe #2613 for ge	c 899	14	10.6	546	22	AAK92496	Human foetal liver
c 827	14	10.6	472	22	AAI34032	Probe #2718 used t	c 900	14	10.6	554	22	ABA60439	Human brain expres
c 828	14	10.6	472	22	AAI02589	Probe #2580 used t	c 901	14	10.6	554	22	AAK80718	Human foetal liver
c 829	14	10.6	472	22	ABS02597	Human genome-deriv	c 902	14	10.6	554	22	AAK34604	Human bone marrow
c 830	14	10.6	473	21	AAZ61808	CDNA encoding muri	c 903	14	10.6	554	22	AAI40323	Probe #9009 used t
c 831	14	10.6	473	22	AAC99741	Skin cell cDNA, SE	c 904	14	10.6	555	24	ABK53445	Human eosinophil-m
c 832	14	10.6	473	22	ABL34893	Murine cDNA isolat	c 905	14	10.6	559	22	AAH97866	Murine 7-transmemb
c 833	14	10.6	474	24	ABL82270	Human ovarian can	c 906	14	10.6	560	22	ABA61635	Human foetal liver
c 834	14	10.6	475	23	ABV35199	Human prostate exp	c 907	14	10.6	560	22	ABA29295	Probe #7761 for ge
c 835	14	10.6	476	22	ABA26701	Probe #5167 for ge	c 908	14	10.6	560	22	AAK09935	Human brain expres
836	14	10.6	476	22	AAH71693	Human cervical can	c 909	14	10.6	560	22	AAK35830	Human bone marrow
837	14	10.6	476	22	AAH72522	Human cervical can	c 910	14	10.6	560	22	AAI41547	Probe #10233 used
838	14	10.6	476	22	AAS31050	Human diagnostic a	c 911	14	10.6	564	24	ABS10064	Human genome-deriv
c 839	14	10.6	477	24	ABK63335	Rat sequence diffe	c 912	14	10.6	564	22	ABA31236	Murine 7-transmemb
840	14	10.6	480	23	ABV03982	Human prostate exp	c 913	14	10.6	564	22	AAI19068	Probe #9702 for ge
841	14	10.6	480	22	ABA2697	Human breast cell	c 914	14	10.6	564	22	AAI44217	Probe #9001 for ge
842	14	10.6	486	22	ABA53120	Human foetal liver	c 915	14	10.6	564	22	ABS12321	Probe #12903 used
843	14	10.6	486	22	ABA22894	Probe #1360 for ge	c 916	14	10.6	566	22	AAS60420	Human genome-deriv
844	14	10.6	486	22	AAK01370	Human brain expres	c 917	14	10.6	567	24	ABN60925	Human cancer relat
845	14	10.6	486	22	AAK26825	Human bone marrow	c 918	14	10.6	568	22	ABA62343	Human foetal liver
846	14	10.6	486	22	AAI11456	Probe #1389 for ge	c 919	14	10.6	568	22	ABA29684	Probe #8150 for ge
847	14	10.6	486	22	AAI32727	Probe #1413 used t	c 920	14	10.6	568	22	AAI10186	Human breast cance
848	14	10.6	486	22	AAI01372	Probe #1363 used t	c 921	14	10.6	568	22	AAK10677	Human brain expres
849	14	10.6	486	24	ABS01423	Human genome-deriv	c 922	14	10.6	568	22	AAK36563	Human bone marrow
850	14	10.6	497	17	AAC73959	Chromosome 6q27 va	c 923	14	10.6	568	22	AAI17419	Probe #7352 for ge
c 851	14	10.6	497	21	AAC75204	Human ORFX ORF759	c 924	14	10.6	568	22	AAI42321	Probe #11007 used
c 852	14	10.6	500	21	AAC95267	Cat flea head and	c 925	14	10.6	568	24	ABS10560	Human genome-deriv
853	14	10.6	500	21	AAZ35046	Mouse EST AA619904	c 926	14	10.6	569	22	ABA59963	Human foetal liver
854	14	10.6	500	22	ABA61624	Human foetal liver	c 927	14	10.6	569	22	ABA28379	Probe #6845 for ge
855	14	10.6	500	22	ABA29287	Probe #7753 for ge	c 928	14	10.6	569	22	AAK08223	Human brain expres
856	14	10.6	500	22	AAK09924	Human brain expres	c 929	14	10.6	569	22	AAK34112	Human bone marrow
857	14	10.6	500	22	AAK35819	Human bone marrow	c 930	14	10.6	569	22	AAK39835	Probe #8521 used t
858	14	10.6	500	22	AAI17145	Probe #7078 for ge	c 931	14	10.6	569	24	ABS08880	Human genome-deriv
859	14	10.6	500	22	AAI41536	Probe #10222 used	c 932	14	10.6	570	22	AAK31073	Human diagnostic a
860	14	10.6	500	24	ABS10054	Human genome-deriv	c 933	14	10.6	570	23	ABV02703	Human prostate exp
861	14	10.6	501	22	ABA12010	Human nervous syst	c 934	14	10.6	571	21	AAK69796	Human ovarian carc
862	14	10.6	503	23	AAS66282	DNA encoding novel	c 935	14	10.6	571	24	ABN72690	Ovarian carcinoma
863	14	10.6	506	21	AAC37038	Arabidopsis thalia	c 936	14	10.6	578	22	AAH07261	Human cDNA clone (
c 864	14	10.6	516	20	AAZ34293	Human EST DNA24256	c 937	14	10.6	580	24	ABN65651	Human cancer relat
c 865	14	10.6	516	21	AAC78581	Human EST DNA24256	c 938	14	10.6	582	23	ABV02797	Human prostate exp
866	14	10.6	518	22	ABA62830	Human foetal liver	c 939	14	10.6	585	22	ABA63060	Human foetal liver
867	14	10.6	518	24	ABS11031	Human genome-deriv	c 940	14	10.6	585	22	ABA30317	Probe #8783 for ge
868	14	10.6	520	22	ABA61592	Human foetal liver	c 941	14	10.6	585	22	AAK11484	Human brain expres
c 869	14	10.6	520	22	AAK09895	Human brain expres	c 942	14	10.6	585	22	AAK37264	Human bone marrow
c 870	14	10.6	521	24	AAS61798	Lung small cell ca	c 943	14	10.6	585	22	AAI18093	Probe #8026 for ge
871	14	10.6	524	24	AAK61761	Lung small cell ca	c 944	14	10.6	585	22	AAI43027	Probe #11783 used
c 872	14	10.6	525	23	ABV11497	Human nervous syst	c 945	14	10.6	585	22	ABS11258	Human genome-deriv
c 873	14	10.6	527	23	ABV55621	Human prostate exp	c 946	14	10.6	586	22	ABA63402	Human foetal liver
874	14	10.6	528	21	AAI14445	Aspergillus oryzae	c 947	14	10.6	586	22	ABA30602	Probe #9068 for ge
875	14	10.6	529	23	ABV58861	Human prostate exp	c 948	14	10.6	586	22	AAK11936	Human brain expres
876	14	10.6	529	23	AAS76160	DNA encoding novel	c 949	14	10.6	586	22	AAK37642	Human bone marrow
877	14	10.6	531	21	AAK69828	Human ovarian carc	c 950	14	10.6	586	22	AAI18397	Probe #8330 for ge
c 878	14	10.6	531	22	ABA60989	Human foetal liver	c 951	14	10.6	586	22	AAI43513	Probe #12199 used
c 879	14	10.6	531	22	ABA28928	Probe #7394 for ge	c 952	14	10.6	586	24	ABS11628	Human genome-deriv
c 880	14	10.6	531	22	AAK09281	Human brain expres	c 953	14	10.6	586	24	ABQ58834	Human colon cancer
c 881	14	10.6	531	22	AAK35170	Human bone marrow	c 954	14	10.6	588	22	AAK63292	Human immune/haema
c 882	14	10.6	531	22	AAI40886	Probe #9572 used t	c 955	14	10.6	589	22	AAK61566	Human foetal liver
c 883	14	10.6	531	24	ABS09609	Human genome-deriv	c 956	14	10.6	589	22	ABA63408	Human foetal liver
c 884	14	10.6	531	24	ABN72722	Ovarian carcinoma	c 957	14	10.6	589	22	ABA30609	Probe #9075 for ge
c 885	14	10.6	532	22	AAI18972	Human breast cance	c 958	14	10.6	589	22	AAK09870	Human brain expres

c 959	14	10.6	589	22	AAK11942
c 960	14	10.6	589	22	AAK35763
c 961	14	10.6	589	22	AAK37648
c 962	14	10.6	589	22	AAI18403
c 963	14	10.6	589	22	AAI14180
c 964	14	10.6	589	22	AAI43519
c 965	14	10.6	589	22	AAI1634
c 966	14	10.6	591	22	AAI1534
c 967	14	10.6	591	22	AAI01169
c 968	14	10.6	591	22	ABV56987
c 969	14	10.6	591	22	ABL96630
c 970	14	10.6	594	21	AAI78045
c 971	14	10.6	594	22	AAI60596
c 972	14	10.6	598	20	AAI39830
c 973	14	10.6	601	24	ABQ66389
c 974	14	10.6	603	23	AAI74748
c 975	14	10.6	603	24	ABK62345
c 976	14	10.6	605	22	AAH97868
c 977	14	10.6	611	23	ABV33017
c 978	14	10.6	611	23	ABV41942
c 979	14	10.6	615	22	AAH05221
c 980	14	10.6	616	24	ABL52680
c 981	14	10.6	620	24	ABK30876
c 982	14	10.6	621	22	AAH87655
c 983	14	10.6	623	21	AAH11677
c 984	14	10.6	623	24	ABN65844
c 985	14	10.6	625	22	AAK88517
c 986	14	10.6	625	24	ABL49421
c 987	14	10.6	625	24	ABK63373
c 988	14	10.6	625	24	ABK30607
c 989	14	10.6	625	24	ABK30964
c 990	14	10.6	626	23	ABV34271
c 991	14	10.6	632	22	AAI15950
c 992	14	10.6	632	22	AAI01348
c 993	14	10.6	632	23	ABL96801
c 994	14	10.6	633	22	AAI60356
c 995	14	10.6	636	23	AAI76366
c 996	14	10.6	641	22	AAH32624
c 997	14	10.6	641	24	ABQ56205
c 998	14	10.6	645	22	AAI41085
c 999	14	10.6	645	24	ABN69799
1000	14	10.6	649	21	AAI14615

ALIGNMENTS

RESULT 1
 ID ABA92649 standard; cDNA: 147 BP.
 XX ABA92649;
 AC ABA92649;
 DT 25-MAR-2002 (first entry)
 XX Human NOV2a encoding polynucleotide SEQ ID NO:3.
 DE
 XX Human; NOVX; cystostatic; antidiabetic; anorectic; antibacterial;
 KW fungicide; virucide; protozoacide; analgesic; antiparkinsonian;
 KW antiasthmatic; hypotensive; osteopathic; antiinflammatory; antiulcer;
 KW neuroprotective; cardiant; antiallergic; antidepressant; nootropic;
 KW anticonvulsant; neuroleptic; gene therapy; vaccine; immune disease;
 KW developmental disease; taste and scent detectability disorder; infection;
 KW Burkitt's lymphoma; signal transduction pathway disorder; pain; cancer;
 KW retinal disease; feeding disorder; asthma; Parkinson's disease; ulcer;
 KW noninsulin-dependent diabetes mellitus; acute heart failure; hypotension;
 KW hypertension; urinary retention; osteoporosis; Crohn's disease; allergy;
 KW multiple sclerosis; angina pectoris; myocardial infarction; delirium;
 KW benign prostatic hypertrophy; manic depression; dementia; dyskinesia;
 KW severe metal retardation; Huntington's disease; gene; chromosome 9;
 KW Gilles de la Tourette syndrome; ss.
 XX Homo sapiens.
 OS

XX	Key	Location/Qualifiers
XX	CDS	6...137
XX		/tag= a
XX		/product= "NOV2a"
XX		WO200190155-A2.
XX		29-NOV-2001.
XX		24-MAY-2001; 2001WO-US17073.
XX		24-MAY-2000; 2000US-206679P.
XX		24-MAY-2000; 2000US-206688P.
XX		24-MAY-2000; 2000US-206829P.
XX		24-MAY-2000; 2000US-207748P.
XX		30-MAY-2000; 2000US-207798P.
XX		30-MAY-2000; 2000US-208263P.
XX		02-JUN-2000; 2000US-208831P.
XX		05-JUN-2000; 2000US-209451P.
XX		07-JUN-2000; 2000US-210060P.
XX		20-JUL-2000; 2000US-219507P.
XX		26-JUL-2000; 2000US-221337P.
XX		31-JUL-2000; 2000US-221927P.
XX		19-JAN-2001; 2001US-263135P.
XX		24-JAN-2001; 2001US-263688P.
XX		23-MAY-2001; 2001US-0863776.
XX		(CURA-) CURAGEN CORP.
XX		Spytek KA, Majumder K, Tchernev VT, Mishra V, Padigaru M;
XX		Spaderna SK, Shenoy S, Rastelli L, Li L, Taupier RJ, Gangolli E;
XX		WPI; 2002-106174/14.
XX		P-PSDB; ABB05034.
XX		Novel polypeptide, useful for treating pain, cancer, urinary retention,
XX		osteoporosis, Crohn's disease, diabetes, acute heart failure, dementia,
XX		asthma, ulcer, allergy and Huntington's disease, comprises isolated
XX		polypeptide NOVX -
XX		Claim 8; page 20; 266pp; English.
XX		The present invention describes human NOVX proteins (NOV1 to NOV9). NOVX
XX		proteins (I) have cytostatic, antidiabetic, anorectic, antibacterial,
XX		fungicide, virucide, protozoacide, analgesic, antiparkinsonian,
XX		antiasthmatic, hypotensive, osteopathic, antiinflammatory, antiulcer,
XX		neuroprotective, cardiant, antiallergic, antidepressant, nootropic,
XX		anticonvulsant and neuroleptic activities. (I) and polynucleotides (II)
XX		can be used in gene therapy and vaccine production. (I) and (II) can be
XX		used for treating or preventing a NOVX-associated disorder such as
XX		cardiomyopathy, atherosclerosis and diabetes in a human, where the
XX		disorder is related to cell signal processing and metabolic pathway
XX		modulation, in a subject, preferably human. (I) and (II) can be used for
XX		diagnosing, preventing or treating developmental diseases, immune
XX		diseases, taste and scent detectability disorder, retinal diseases including those
XX		involving photoreception, cell growth rate disorders, feeding disorders,
XX		noninsulin-dependent diabetes mellitus, infections, pain, cancer, asthma,
XX		Parkinson's disease, acute heart failure, hypotension, hypertension,
XX		urinary retention, osteoporosis, Crohn's disease, multiple sclerosis,
XX		Albright Hereditary Osteodystrophy, angina pectoris, ulcers, myocardial
XX		infarction, allergies, benign prostatic hypertrophy, manic depression,
XX		delirium, dementia, severe metal retardation and dyskinesias, such as
XX		Huntington's disease or Gilles de la Tourette syndrome. The present
XX		sequence encodes human NOV2a, which is isolated to chromosome 9.
XX		Sequence 147 BP; 56 A; 30 C; 41 G; 20 T; 0 other;
XX		Query Match 100.0%; Score 132; DB 24; Length 147;
XX		Best Local Similarity 100.0%; Pred. No. 2.9e-56;
XX		Matches 132; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ATGGCACAAACTAGACTGGAGAAATTCAGCTTGATAAGGCCAAGCTGAAGGCC 60
DB 6 ATGGCACAAACTAGACTGGAGAAATTCAGCTTGATAAGGCCAAGCTGAAGGCC 65
OY 61 ACAGAGATGCAGAGAACTCTGATGACCAAGAGACACAGAGCAGGAGAGTGGAGT 120
DB 66 ACAGAGATGCAGAGAACTCTGATGACCAAGAGACACAGAGCAGGAGAGTGGAGT 125
OY 121 GAAATTTCTCTGA 132
DB 126 GAAATTTCTCTGA 137

RESULT 2

ABA92650

ID ABA92650 standard; cDNA; 147 BP.

XX AC ABA92650;

XX 25-MAR-2002 (first entry)

DE Human NOV2b encoding polynucleotide SEQ ID NO:5.

XX Human; NOVX; cytostatic; antidiabetic; anorectic; antibacterial;
KW fungicide; virucide; protozoacide; analgesic; antiparkinsonian;
KW antiasthmatic; hypotensive; osteopathic; antiinflammatory; antiulcer;
KW neuroprotective; cardiant; antiallergic; antidepressant; nootropic;
KW anticonvulsant; neuroleptic; gene therapy; vaccine; immune disease;
KW developmental disease; taste and scent detectability disorder; infection;
KW Burkitt's lymphoma; signal transduction pathway disorder; pain; cancer;
KW retinal disease; feeding disorder; asthma; Parkinson's disease; ulcer;
KW noninsulin-dependent diabetes mellitus; acute heart failure; hypotension;
KW hypertension; urinary retention; osteoporosis; Crohn's disease; allergy;
KW multiple sclerosis; angina pectoris; myocardial infarction; delirium;
KW benign prostatic hypertrophy; manic depression; dementia; dyskinesia;
KW severe mental retardation; Huntington's disease; gene;
KW Gilles de la Tourette syndrome; ss.

OS Homo sapiens.

XX Key Location/Qualifiers

FT CDS 6..137

FT /*tag= a

FT /product= "NOV2b"

XX WO200190155-A2.

XX 29-NOV-2001.

XX 24-MAY-2001; 2001WO-US17073.

XX 24-MAY-2000; 2000US-206679P.

XX 24-MAY-2000; 2000US-206688P.

XX 24-MAY-2000; 2000US-206829P.

XX 30-MAY-2000; 2000US-207748P.

XX 30-MAY-2000; 2000US-207798P.

XX 31-MAY-2000; 2000US-208263P.

XX 02-JUN-2000; 2000US-208831P.

XX 05-JUN-2000; 2000US-209451P.

XX 07-JUN-2000; 2000US-210060P.

XX 20-JUL-2000; 2000US-219507P.

XX 26-JUL-2000; 2000US-221337P.

XX 31-JUL-2000; 2000US-221927P.

XX 19-JAN-2001; 2001US-263135P.

XX 24-JAN-2001; 2001US-263688P.

XX 24-JAN-2001; 2001US-263694P.

XX 23-MAY-2001; 2001US-0863776.

XX (CURA-) CURAGEN CORP.

XX Spytek KA, Majumder K, Tchernev VT, Mishra V, Padigaru M;

PI Spaderna SK, Shenoy S, Rastelli L, Li L, Taupier RJ, Gangolli E;

XX WPI: 2002-106174/14.
DR P-PSDB; ABB05035.

XX Novel polypeptide, useful for treating pain, cancer, urinary retention,
PT osteoporosis, Crohn's disease, diabetes, acute heart failure, dementia,
PT asthma, ulcer, allergy and Huntington's disease, comprises isolated
PT polypeptide NOVX -

XX Claim 8; page 21; 266pp; English.

XX The present invention describes human NOVX proteins (NOV1 to NOV9). NOVX
CC proteins (I) have cytostatic, antidiabetic, anorectic, antibacterial,
CC fungicide, virucide, protozoacide, analgesic, antiparkinsonian,
CC antiasthmatic, hypotensive, osteopathic, antiinflammatory, antiulcer,
CC neuroprotective, cardiant, antiallergic, antidepressant, nootropic,
CC anticonvulsant and neuroleptic activities. (I) and polynucleotides (II)
CC can be used in gene therapy and vaccine production. (I) and (II) can be
CC used for treating or preventing a NOVX-associated disorder such as
CC cardiomyopathy, atherosclerosis and diabetes in a human, where the
CC disorder is related to cell signal processing and metabolic path way
CC modulation, in a subject, preferably human. (I) and (II) can be used for
CC diagnosing, preventing or treating developmental diseases, immune
CC diseases, taste and scent detectability disorder, Burkitt's lymphoma,
CC signal transduction pathway disorders, retinal diseases including those
CC involving photoreception, cell growth rate disorders, feeding disorders,
CC noninsulin-dependent diabetes mellitus, infections, pain, cancer, asthma,
CC Parkinson's disease, acute heart failure, hypotension, hypertension,
CC urinary retention, osteoporosis, Crohn's disease, multiple sclerosis,
CC Albright Hereditary Osteodystrophy, angina pectoris, ulcers, myocardial
CC infarction, allergies, benign prostatic hypertrophy, manic depression,
CC delirium, dementia, severe mental retardation and dyskinesias, such as
CC Huntington's disease or Gilles de la Tourette syndrome. The present
CC sequence encodes human NOV2b, from the present invention.

XX Sequence 147 BP; 56 A; 30 C; 41 G; 20 T; 0 other;

Query Match 100.0%; Score 132; DB 24; Length 147;
Best Local Similarity 100.0%; Pred. No. 2.9e-56;
Matches 132; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ATGGCACAAACTAGACTGGAGAAATTCAGCTTGATAAGGCCAAGCTGAAGGCC 60

DB 6 ATGGCACAAACTAGACTGGAGAAATTCAGCTTGATAAGGCCAAGCTGAAGGCC 65

OY 61 ACAGAGATGCAGAGAACTCTGATGACCAAGAGACACAGAGCAGGAGAGTGGAGT 120

DB 66 ACAGAGATGCAGAGAACTCTGATGACCAAGAGACACAGAGCAGGAGAGTGGAGT 125

OY 121 GAAATTTCTCTGA 132

DB 126 GAAATTTCTCTGA 137

RESULT 3

ABA92651

ID ABA92651 standard; cDNA; 147 BP.

XX ABA92651;

XX 25-MAR-2002 (first entry)

DE Human NOV2c encoding polynucleotide SEQ ID NO:7.

XX Human; NOVX; cytostatic; antidiabetic; anorectic; antibacterial;
KW fungicide; virucide; protozoacide; analgesic; antiparkinsonian;
KW antiasthmatic; hypotensive; osteopathic; antiinflammatory; antiulcer;
KW neuroprotective; cardiant; antiallergic; antidepressant; nootropic;
KW anticonvulsant; neuroleptic; gene therapy; vaccine; immune disease;
KW developmental disease; taste and scent detectability disorder; infection;
KW Burkitt's lymphoma; signal transduction pathway disorder; pain; cancer;
KW retinal disease; feeding disorder; asthma; Parkinson's disease; ulcer;
KW noninsulin-dependent diabetes mellitus; acute heart failure; hypotension;

CC sequence encodes human NOV2c, from the present invention.
XX Sequence 147 BP; 56 A; 30 C; 41 G; 20 T; 0 other;
SQ
Query Match 100.0%; Score 132; DB 24; Length 147;
Best Local Similarity 100.0%; Pred. NO. 2.9e-56;
Matches 132; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGCGACACAACTAGACCTGGAAGAATTCGACCTTGGGATAAGCCCAAGCTGAAGGCC 60
DB 6 ATGCGACACAACTAGACCTGGAAGAATTCGACCTTGGGATAAGCCCAAGCTGAAGGCC 65
QY 61 ACAGAGATGCGAGAGAACACTCTGATGACCAAGAGACACACAGAGAGAGAGTGGACT 120
DB 66 ACAGAGATGCGAGAGAACACTCTGATGACCAAGAGACACACAGAGAGAGTGGAGT 125
QY 121 GAAATTTCTTGA 132
DB 126 GAAATTTCTTGA 137
RESULT 4
ABN29874
ID ABN29874 standard; DNA; 65 BP.
XX AC ABN29874;
XX DT 15-JUL-2002 (first entry)
XX DE Rat spliced transcript detection oligonucleotide SEQ ID NO:2622.
XX KW Human; mouse; rat; splice transcript; detection; RNA transcript;
XX KW splice variant; transcriptome; oligonucleotide library; ss.
XX OS Rattus norvegicus.
XX PN WO200210449-A2.
XX PD 07-FEB-2002.
XX PF 20-JUL-2001; 2001WO-IB01903.
XX PR 28-JUL-2000; 2000US-221607P.
XX PR 02-MAY-2001; 2001US-287724P.
XX PA (COMP-) COMPUGEN INC.
XX PI Shoshan A, Wasserman A, Mintz E, Mintz L, Faigler S;
XX DR WPI; 2002-257383/30.
XX PT New oligonucleotide libraries comprising oligonucleotides which
PT selectively hybridize to mRNAs transcribed from a transcription unit of
PT a genome, useful for detecting tissue-, pathology-, and
XX developmental-specific genes
PS Example 1; SEQ ID 2622; 47pp; English.
XX The present invention describes oligonucleotide libraries for detecting
CC messenger RNAs that populate a (sub-)transcriptome, where the
CC (sub-)transcriptome comprises messenger RNAs transcribed from multiple
CC transcription units that populate a genome. The library comprises
CC several oligonucleotides, each capable of hybridising selectively to a
CC set of messenger RNAs transcribed from a given transcription unit of
CC the genome, which encodes one or more messenger RNA splice variants.
CC The oligonucleotide libraries are useful for detecting mRNAs from a
CC biological sample, in expression profiling studies, in qualitatively or
CC quantitatively characterising the corresponding transcriptome, and in
CC detecting RNA transcripts and splice variants of human or animal
CC transcriptomes. The libraries may also be used as specialised mini
CC libraries to detect transcripts of a sub-transcriptome under a
CC particular biological or pathological state, and so allowing the
CC detection of tissue- and pathology-specific genes such as those genes

KW hypertension; urinary retention; osteoporosis; Crohn's disease; allergy;
KW multiple sclerosis; angina pectoris; myocardial infarction; delirium;
KW benign prostatic hypertrophy; manic depression; dementia; dyskinesia;
KW severe metal retardation; Huntington's disease; gene;
KW Gilles de la Tourette syndrome; Huntington, ss.
XX
XX Homo sapiens.
XX Key Location/Qualifiers
XX CDS 6..137
XX /*tag= a "NOV2c"
XX /product= "NOV2c"
XX WO200190155-A2.
XX 29-NOV-2001.
XX 24-MAY-2001; 2001WO-US17073.
XX 24-MAY-2000; 2000US-206679P.
XX 24-MAY-2000; 2000US-206688P.
XX 24-MAY-2000; 2000US-206829P.
XX 30-MAY-2000; 2000US-207748P.
XX 30-MAY-2000; 2000US-207798P.
XX 31-MAY-2000; 2000US-208263P.
XX 02-JUN-2000; 2000US-208831P.
XX 05-JUN-2000; 2000US-209451P.
XX 07-JUN-2000; 2000US-210060P.
XX 20-JUL-2000; 2000US-219507P.
XX 26-JUL-2000; 2000US-221337P.
XX 31-JUL-2000; 2000US-221927P.
XX 19-JAN-2001; 2001US-263135P.
XX 24-JAN-2001; 2001US-263688P.
XX 24-JAN-2001; 2001US-263694P.
XX 23-MAY-2001; 2001US-0863776.
XX (CURA-) CURAGEN CORP.
XX Spytek KA, Majumder K, Tchernev VT, Mishra V, Padigar M;
XX Spaderna SK, Shenoy S, Rastelli L, Li L, Taupier RJ, Gangolli E;
XX WPI; 2002-106174/14.
XX P-PSDB; ABB05036.
XX Novel polypeptide, useful for treating pain, cancer, urinary retention,
XX osteoporosis, Crohn's disease, diabetes, acute heart failure, dementia,
XX asthma, ulcer, allergy and Huntington's disease, comprises isolated
XX polypeptide NOVX -
XX Claim 8; page 23; 266pp; English.
XX The present invention describes human NOVX proteins (NOV1 to NOV9). NOVX
XX proteins (I) have cytostatic, antidiabetic, anorectic, antibacterial,
XX fungicide, virucide, protozoacide, analgesic, antiparkinsonian,
XX antiasthmatic, hypotensive, osteopathic, antiinflammatory, antitumor,
XX neuroprotective, cardiant, antiallergic, antidepressant, nootropic,
XX anticonvulsant and neuroleptic activities. (I) and polynucleotides (II)
XX can be used in gene therapy and vaccine production. (I) and (II) can be
XX used for treating or preventing a NOVX-associated disorder such as
XX cardiomyopathy, atherosclerosis and diabetes in a human, where the
XX disorder is related to cell signal processing and metabolic path way
XX modulation, in a subject, preferably human. (I) and (II) can be used for
XX diagnosing, preventing or treating developmental diseases, immune
XX diseases, taste and scent detectability disorder, Burkitt's lymphoma,
XX signal transduction pathway disorders, retinal diseases including those
XX involving photoreception, cell growth rate disorders, feeding disorders,
XX noninsulin-dependent diabetes mellitus, infections, pain, cancer, asthma,
XX Parkinson's disease, acute heart failure, hypotension, hypertension,
XX urinary retention, osteoporosis, Crohn's disease, multiple sclerosis,
XX Albrit Hereditary Osteodystrophy, angina pectoris, ulcers, myocardial
XX infarction, allergies, benign prostatic hypertrophy, manic depression,
XX delirium, dementia, severe metal retardation and dyskinesias, such as
XX Huntington's disease or Gilles de la Tourette syndrome. The present

CC only expressed in specific tissue under a specific pathological
CC condition; to detect developmental specific genes; and to detect RNA
CC transcripts and splice variants of a transcriptome of a patient suffering
CC from a particular disorder. ABN27253 to ABN59589 represent
CC oligonucleotide sequences from rats, humans and mice, which are used in
CC the amplification of the present invention.
CC N.B. The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
SQ Sequence 65 BP; 24 A; 13 C; 20 G; 8 T; 0 other;

Query Match 13.6%; Score 18; DB 24; Length 65;
Best Local Similarity 100.0%; Pred. No. 23;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 40 GATAGGCCCAAGCTGAAG 57
DB. 47 GATAGGCCCAAGCTGAAG 64
|||||
RESULT 5
ABV08667/c
ID ABV08667 standard; cDNA; 203 BP.
XX AC
XX ABV08667;
XX DT 13-SEP-2002 (first entry)
XX DE Human prostate expression marker cDNA 8658.
XX KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
XX KW pharmacogenomic marker; gene; ss.
XX OS Homo sapiens.
XX PN WO200160860-A2.
XX PD 23-AUG-2001.
XX PF 20-FEB-2001; 2001WO-US05171.
XX PR 17-FEB-2000; 2000US-183319P.
XX PR 16-MAR-2000; 2000US-189862P.
XX PR 25-MAY-2000; 2000US-207454P.
XX PR 09-JUN-2000; 2000US-211314P.
XX PR 18-JUL-2000; 2000US-219007P.
XX PR 13-DEC-2000; 2000US-255281P.
XX PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX PI Schlegel R, Endege WO, Monahan JE;
XX WPI; 2001-662795/76.
XX DR
XX PT Novel isolated nucleic acid molecule associated with cancerous state of
XX prostate cells and correlating with presence of prostate cancer, useful
XX for detecting presence of prostate cancer, stage of prostate cancer.
XX PS
XX Claim 1; Page 1368; 11750pp; English.
XX
CC The invention relates to an isolated nucleic acid molecule (I) comprising
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
CC specification or its complement. (I) is useful for:
CC (a) assessing whether a patient is afflicted with prostate cancer;
CC (b) monitoring the progression of prostate cancer in a patient;
CC (c) assessing the efficacy of a test compound to inhibit prostate
CC cancer in a patient;
CC (d) assessing the efficacy of a therapy for inhibiting prostate cancer
CC in a patient;
CC (e) selecting a composition for inhibiting prostate cancer in a patient;
CC (f) assessing the prostate cell carcinogenic potential of a compound;
CC (g) determining whether prostate cancer has metastasized in a patient;

CC (h) assessing the aggressiveness or indolence of prostate cancer in a
CC patient;
CC (I) is also useful as a pharmacodynamic or pharmacogenomic marker.
XX
SQ Sequence 203 BP; 30 A; 65 C; 50 G; 58 T; 0 other;

Query Match 13.6%; Score 18; DB 23; Length 203;
Best Local Similarity 100.0%; Pred. No. 23;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 40 GATAGGCCCAAGCTGAAG 57
DB. 73 GATAGGCCCAAGCTGAAG 56
|||||
RESULT 6
ABK55047
ID ABK55047 standard; cDNA; 243 BP.
XX AC
XX ABK55047;
XX DT 18-JUN-2002 (first entry)
XX DE Human colon cancer-associated cDNA, SEQ ID NO 517.
XX KW Human; colon cancer; immunogenic; vaccine; tumour; gene; ss.
XX OS Homo sapiens.
XX PN WO200212280-A2.
XX PD 14-FEB-2002.
XX PF 30-JUL-2001; 2001WO-US23826.
XX PR 03-AUG-2000; 2000US-223265P.
XX PR 02-OCT-2000; 2000US-237406P.
XX PR 20-MAR-2001; 2001US-277495P.
XX PR 03-JUL-2001; 2001US-302702P.
XX PA (CORI-) CORIXA CORP.
XX PI Pyle RA, Xu J, Secríst H;
XX WPI; 2002-257462/30.
XX DR
XX PT Novel polynucleotide encoding colon tumour polypeptides, useful as
XX vaccines for treating colon cancers -
XX PS
XX Claim 1; Page 297; 425pp; English.
XX
CC The invention relates to isolated polynucleotides (I) encoding colon
CC tumour polypeptides (II). (I) is useful for stimulating an immune
CC response in a patient and treating colon cancer in a patient.
CC Oligonucleotides derived from (I) are useful for determining the presence
CC of cancer in a patient. (I) and (II) are useful in pharmaceutical
CC compositions, e.g. vaccines, and other compositions for the diagnosis
CC and treatment of colon cancer. A composition comprising a first component
CC selected from physiologically acceptable carriers and immunostimulants,
CC and an antigen-presenting cell expressing (II) is useful for inhibiting
CC development of cancer in a patient. (I) is useful in the design and
CC preparation of ribozyme molecules for inhibiting expression of tumour
CC polypeptides and (I). ABK54531-ABK55464 represent human colon cancer cDNA
CC sequences of the invention.
XX
SQ Sequence 243 BP; 73 A; 64 C; 68 G; 37 T; 1 other;

Query Match 13.6%; Score 18; DB 24; Length 243;
Best Local Similarity 100.0%; Pred. No. 23;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 40 GATAGGCCCAAGCTGAAG 57
|||||

Db	102	GATAAGGCCAAGCTGAAG	119
RESULT 7			
AAC000909			
ID	AAC00909	standard; cDNA; 283 BP.	
XX	XX		
AC	AC		
AC	AAC00909;		
XX	XX		
DT	DT		
XX	XX	06-OCT-2000 (first entry)	
DE	DE		
XX	XX	Human secreted protein 5' EST, SEQ ID NO: 907.	
XX	XX		
XX	XX	Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;	
KW	KW	gene therapy; chromosome mapping; ss. .	
XX	XX		
XX	XX		
OS	OS	Homo sapiens.	
XX	XX		
XX	XX	EP1033401-A2.	
PN	PN		
XX	XX	06-SEP-2000.	
XX	XX		
XX	XX	21-FEB-2000; 2000EP-0200610.	
PF	PF		
XX	XX	26-FEB-1999; 99US-0122487.	
PR	PR		
XX	XX		
XX	XX	(GEST) GENSET.	
PA	PA		
XX	XX	Dumas Milne Edwards J, Duclert A, Giordano J;	
PI	PI		
XX	XX	WPI; 2000-500381/45.	
DR	DR	P-FSDB; AAG00903.	
XX	XX		
PT	PT	New nucleic acid that is a 5' expressed sequence tag (5' EST) for	
PT	PT	obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for	
PT	PT	diagnostic, forensic, gene therapy and chromosome mapping procedures -	
XX	XX		
PS	PS	Claim 1; SEQ ID 907; 71pp + CD-ROM; English.	
XX	XX		
CC	CC	The present sequence is one of a large number of 5' ESTs derived from	
CC	CC	mRNAs encoding secreted proteins. An ORF has been identified within the	
CC	CC	sequence. The 5' ESTs were prepared from total human RNAs or polyA+ RNAs	
CC	CC	derived from 30 different tissues. EST sequences usually correspond	
CC	CC	mainly to the 3' untranslated region (UTR) of the mRNA because they are	
CC	CC	often obtained from oligo-dT primed cDNA libraries. Such ESTs are not	
CC	CC	well suited for isolating cDNA sequences derived from the 5' ends of	
CC	CC	mRNAs and even in those cases where longer cDNA sequences have been	
CC	CC	obtained, the full 5' UTR is rarely included. 5' ESTs are derived from	
CC	CC	mRNAs with intact 5' ends and can therefore be used to obtain full length	
CC	CC	cDNAs and genomic DNAs. 5' ESTs are also used in diagnostic, forensic,	
CC	CC	gene therapy and chromosome mapping procedures. They are used to obtain	
CC	CC	upstream regulatory sequences and to design expression and secretion	
XX	XX	vectors.	
XX	XX		
SQ	SQ	Sequence 283 BP; 71 A; 87 C; 78 G; 47 T; 0 other;	
		Query Match 13.6%; Score 18; DB 21; Length 283;	
		Best Local Similarity 100.0%; Pred. No. 23;	
		Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY	QY	40 GATAAGGCCAAGCTGAAG 57	
Db	Db	147 GATAAGGCCAAGCTGAAG 164	
RESULT 8			
ABK54754			
ID	ABK54754	standard; cDNA; 353 BP.	
XX	XX		
XX	XX	ABK54754;	
XX	XX		
DT	DT	18-JUN-2002 (first entry)	
XX	XX		

PR 03-AUG-2000; 2000US-223265P.
PR 02-OCT-2000; 2000US-237406P.
PR 20-MAR-2001; 2001US-277495P.
PR 03-JUL-2001; 2001US-302702P.
PA (CORI-) CORIXA CORP.
XX
XX
PI Pyle RA, Xu J, Secrlist H;
XX
XX WPI; 2002-257462/30.
DR
XX
XX Novel polynucleotide encoding colon tumour polypeptides, useful as
PT vaccines for treating colon cancers -
XX
XX
PS Claim 1; Page 209; 425pp; English.
XX
XX The invention relates to isolated polynucleotides (I) encoding colon
CC tumour polypeptides (II). (I) is useful for stimulating an immune
CC response in a patient and treating colon cancer in a patient.
CC Oligonucleotides derived from (I) are useful for determining the presence
CC of cancer in a patient. (I) and (II) are useful in pharmaceutical
CC compositions, e.g. vaccines, and other compositions for the diagnosis
CC and treatment of colon cancer. A composition comprising a first component
CC selected from physiologically acceptable carriers and immunostimulants,
CC and an antigen-presenting cell expressing (II) is useful for inhibiting
CC development of cancer in a patient. (I) is useful in the design and
CC preparation of ribozyme molecules for inhibiting expression of tumour
CC polypeptides and (I). ABK54531-ABK55464 represent human colon cancer cDNA
CC sequences of the invention.
XX
XX Sequence 395 BP; 105 A; 113 C; 115 G; 62 T; 0 other;
SQ
Query Match 13.6%; Score 18; DB 24; Length 395;
Best Local Similarity 100.0%; Pred. No. 23;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 40 GATAAGGCCAAGCTGAAG 57
|||||
DB 111 GATAAGGCCAAGCTGAAG 128
RESULT 10
ABV38568/c
ID ABV38568 standard; cDNA; 399 BP.
XX
AC ABV38568;
XX
DT 16-SEP-2002 (first entry)
XX
DE Human prostate expression marker cDNA 38559.
XX
XX Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
KW pharmacogenomic marker; gene; ss.
XX
XX Homo sapiens.
OS
XX WO200160860-A2.
PN
XX 23-AUG-2001.
PD
XX
XX 20-FEB-2001; 2001WO-US05171.
PF
XX
XX 17-FEB-2000; 2000US-183319P.
PR 16-MAR-2000; 2000US-189862P.
PR 25-MAY-2000; 2000US-207454P.
PR 09-JUN-2000; 2000US-211314P.
PR 18-JUL-2000; 2000US-219007P.
PR 13-DEC-2000; 2000US-255281P.
XX
XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
PA
XX
PI Schlegel R, Endege WO, Monahan JE;
XX

DR WPI; 2001-662795/76.
XX
XX Novel isolated nucleic acid molecule associated with cancerous state of
PT prostate cells and correlating with presence of prostate cancer, useful
PT for detecting presence of prostate cancer, stage of prostate cancer -
XX
XX
PS Claim 1; Page 7850; 11750pp; English.
XX
XX The invention relates to an isolated nucleic acid molecule (I) comprising
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
CC specification or its complement. (I) is useful for:
CC (a) assessing whether a patient is afflicted with prostate cancer;
CC (b) monitoring the progression of prostate cancer in a patient;
CC (c) assessing the efficacy of a test compound to inhibit prostate
CC cancer in a patient;
CC (d) assessing the efficacy of a therapy for inhibiting prostate cancer
CC in a patient;
CC (e) selecting a composition for inhibiting prostate cancer in a patient;
CC (f) assessing the prostate cell carcinogenic potential of a compound;
CC (g) determining whether prostate cancer has metastasized in a patient;
CC (h) assessing the aggressiveness or indolence of prostate cancer in a
CC patient;
CC (I) is also useful as a pharmacodynamic or pharmacogenomic marker.
XX
XX Sequence 399 BP; 70 A; 116 C; 100 G; 113 T; 0 other;
SQ
Query Match 13.6%; Score 18; DB 23; Length 399;
Best Local Similarity 100.0%; Pred. No. 23;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 40 GATAAGGCCAAGCTGAAG 57
|||||
DB 78 GATAAGGCCAAGCTGAAG 61
RESULT 11
ABK55033
ID ABK55033 standard; cDNA; 428 BP.
XX
XX
AC ABK55033;
XX
DT 18-JUN-2002 (first entry)
XX
DE Human colon cancer-associated cDNA, SEQ ID No 503.
XX
XX Human; colon cancer; immunogenic; vaccine; tumour; gene; ss.
KW
XX
XX Homo sapiens.
OS
XX WO200212280-A2.
PN
XX
XX 14-FEB-2002.
PD
XX
XX 30-JUL-2001; 2001WO-US23826.
PF
XX
XX 03-AUG-2000; 2000US-223265P.
PR 02-OCT-2000; 2000US-237406P.
PR 20-MAR-2001; 2001US-277495P.
PR 03-JUL-2001; 2001US-302702P.
XX
XX (CORI-) CORIXA CORP.
PA
XX
XX Pyle RA, Xu J, Secrlist H;
PI
XX
XX WPI; 2002-257462/30.
DR
XX
XX Novel polynucleotide encoding colon tumour polypeptides, useful as
PT vaccines for treating colon cancers -
XX
XX
PS Claim 1; Page 294; 425pp; English.
XX
XX The invention relates to isolated polynucleotides (I) encoding colon
CC tumour polypeptides (II). (I) is useful for stimulating an immune

CC response in a patient and treating colon cancer in a patient.
 CC Oligonucleotides derived from (I) are useful for determining the presence
 CC of cancer in a patient. (I) and (II) are useful in pharmaceutical
 CC compositions, e.g. vaccines, and other compositions for the diagnosis
 CC and treatment of colon cancer. A composition comprising a first component
 CC selected from physiologically acceptable carriers and immunostimulants,
 CC and an antigen-presenting cell expressing (II) is useful for inhibiting
 CC development of cancer in a patient. (I) is useful in the design and
 CC preparation of ribozyme molecules for inhibiting expression of tumour
 CC polypeptides and (I). ABK54531-ABK55464 represent human colon cancer cDNA
 CC sequences of the invention.

XX SQ Sequence 428 BP; 124 A; 114 C; 115 G; 75 T; 0 other;

Query Match 13.6%; Score 18; DB 24; Length 428;
 Best Local Similarity 100.0%; Pred. No. 23;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 40 GATAAGGCCAAGCTGAAG 57
 Db 89 GATAAGGCCAAGCTGAAG 106
 |||||

RESULT 12
 ABK55163
 ID ABK55163 standard; cDNA; 429 BP.

XX AC ABK55163;
 XX DT 18-JUN-2002 (first entry)
 XX DE Human colon cancer-associated cDNA, SEQ ID No 633.
 XX KW Human; colon cancer; immunogenic; vaccine; tumour; gene; ss.
 XX OS Homo sapiens.

XX PN WO200212280-A2.
 XX PD 14-FEB-2002.
 XX PF 30-JUL-2001; 2001WO-US23826.
 XX PR 03-AUG-2000; 2000US-223265P.
 XX PR 02-OCT-2000; 2000US-237406P.
 XX PR 20-MAR-2001; 2001US-277495P.
 XX PR 03-JUL-2001; 2001US-302702P.

XX PA (CORI-) CORIXA CORP.
 XX PI Pyle RA, Xu J, Secretist H;
 XX DR WPI; 2002-257462/30.
 XX PT Novel polynucleotide encoding colon tumour polypeptides, useful as
 XX vaccines for treating colon cancers

XX PS Claim 1; Page 329; 425pp; English.
 XX The invention relates to isolated polynucleotides (I) encoding colon
 CC tumour polypeptides (II). (I) is useful for stimulating an immune
 CC response in a patient and treating colon cancer in a patient.
 CC Oligonucleotides derived from (I) are useful for determining the presence
 CC of cancer in a patient. (I) and (II) are useful in pharmaceutical
 CC compositions, e.g. vaccines, and other compositions for the diagnosis
 CC and treatment of colon cancer. A composition comprising a first component
 CC selected from physiologically acceptable carriers and immunostimulants,
 CC and an antigen-presenting cell expressing (II) is useful for inhibiting
 CC development of cancer in a patient. (I) is useful in the design and
 CC preparation of ribozyme molecules for inhibiting expression of tumour
 CC polypeptides and (I). ABK54531-ABK55464 represent human colon cancer cDNA
 CC sequences of the invention.

XX

SQ Sequence 429 BP; 124 A; 115 C; 115 G; 75 T; 0 other;

Query Match 13.6%; Score 18; DB 24; Length 429;
 Best Local Similarity 100.0%; Pred. No. 23;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 40 GATAAGGCCAAGCTGAAG 57
 Db 90 GATAAGGCCAAGCTGAAG 107
 |||||

RESULT 13
 ABK55157
 ID ABK55157 standard; cDNA; 434 BP.

XX AC ABK55157;
 XX DT 18-JUN-2002 (first entry)
 XX DE Human colon cancer-associated cDNA, SEQ ID No 627.
 XX KW Human; colon cancer; immunogenic; vaccine; tumour; gene; ss.
 XX OS Homo sapiens.

XX PN WO200212280-A2.
 XX PD 14-FEB-2002.
 XX PF 30-JUL-2001; 2001WO-US23826.
 XX PR 03-AUG-2000; 2000US-223265P.
 XX PR 02-OCT-2000; 2000US-237406P.
 XX PR 20-MAR-2001; 2001US-277495P.
 XX PR 03-JUL-2001; 2001US-302702P.
 XX PA (CORI-) CORIXA CORP.

XX PI Pyle RA, Xu J, Secretist H;
 XX DR WPI; 2002-257462/30.
 XX PT Novel polynucleotide encoding colon tumour polypeptides, useful as
 XX vaccines for treating colon cancers

XX PS Claim 1; Page 328; 425pp; English.
 XX The invention relates to isolated polynucleotides (I) encoding colon
 CC tumour polypeptides (II). (I) is useful for stimulating an immune
 CC response in a patient and treating colon cancer in a patient.
 CC Oligonucleotides derived from (I) are useful for determining the presence
 CC of cancer in a patient. (I) and (II) are useful in pharmaceutical
 CC compositions, e.g. vaccines, and other compositions for the diagnosis
 CC and treatment of colon cancer. A composition comprising a first component
 CC selected from physiologically acceptable carriers and immunostimulants,
 CC and an antigen-presenting cell expressing (II) is useful for inhibiting
 CC development of cancer in a patient. (I) is useful in the design and
 CC preparation of ribozyme molecules for inhibiting expression of tumour
 CC polypeptides and (I). ABK54531-ABK55464 represent human colon cancer cDNA
 CC sequences of the invention.

XX SQ Sequence 434 BP; 124 A; 116 C; 117 G; 77 T; 0 other;

Query Match 13.6%; Score 18; DB 24; Length 434;
 Best Local Similarity 100.0%; Pred. No. 23;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 40 GATAAGGCCAAGCTGAAG 57
 Db 93 GATAAGGCCAAGCTGAAG 110
 |||||

RESULT 14

ABK54988
ID ABK54988 standard; cDNA; 438 BP.

XX AC ABK54988;

XX DT 18-JUN-2002 (first entry)

XX DE Human colon cancer-associated cDNA, SEQ ID No 458.

XX KW Human; colon cancer; immunogenic; vaccine; tumour; gene; ss.

XX OS Homo sapiens.

XX PN WO200212280-A2.

XX PD 14-FEB-2002.

XX PF 30-JUL-2001; 2001WO-US23826.

XX PR 03-AUG-2000; 2000US-223265P.

XX PR 02-OCT-2000; 2000US-237406P.

XX PR 20-MAR-2001; 2001US-277495P.

XX PR 03-JUL-2001; 2001US-302702P.

XX PA (CORI-) CORIXA CORP.

XX PI Pyle RA, Xu J, Secretist H;

XX DR WPI; 2002-257462/30.

XX PT Novel polynucleotide encoding colon tumour polypeptides, useful as
PT vaccines for treating colon cancers -

XX PS Claim 1; Page 280-281; 425pp; English.

XX CC The invention relates to isolated polynucleotides (I) encoding colon
CC tumour polypeptides (II). (I) is useful for stimulating an immune
CC response in a patient and treating colon cancer in a patient.
CC Oligonucleotides derived from (I) are useful for determining the presence
CC of cancer in a patient. (I) and (II) are useful in pharmaceutical
CC compositions, e.g. vaccines, and other compositions for the diagnosis
CC and treatment of colon cancer. A composition comprising a first component
CC selected from physiologically acceptable carriers and immunostimulants,
CC and an antigen-presenting cell expressing (II) is useful for inhibiting
CC development of cancer in a patient. (I) is useful in the design and
CC preparation of ribozyme molecules for inhibiting expression of tumour
CC polypeptides and (I). ABK54531-ABK55464 represent human colon cancer cDNA
CC sequences of the invention.

XX SQ Sequence 438 BP; 126 A; 118 C; 117 G; 77 T; 0 other;

Query Match 13.6%; Score 18; DB 24; Length 438;
Best Local Similarity 100.0%; Pred. No. 23;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 40 GATAAGGCCAAGCTGAAG 57

DB 99 GATAAGGCCAAGCTGAAG 116

RESULT 15

ABK55113
ID ABK55113 standard; cDNA; 438 BP.

XX AC ABK55113;

XX DT 18-JUN-2002 (first entry)

XX DE Human colon cancer-associated cDNA, SEQ ID No 583.

XX KW Human; colon cancer; immunogenic; vaccine; tumour; gene; ss.

XX OS Homo sapiens.

XX PN WO200212280-A2.

XX PD 14-FEB-2002.

XX PF 30-JUL-2001; 2001WO-US23826.

XX PR 03-AUG-2000; 2000US-223265P.

XX PR 02-OCT-2000; 2000US-237406P.

XX PR 20-MAR-2001; 2001US-277495P.

XX PR 03-JUL-2001; 2001US-302702P.

XX PA (CORI-) CORIXA CORP.

XX PI Pyle RA, Xu J, Secretist H;

XX DR WPI; 2002-257462/30.

XX PT Novel polynucleotide encoding colon tumour polypeptides, useful as
PT vaccines for treating colon cancers -

XX PS Claim 1; Page 315-316; 425pp; English.

XX CC The invention relates to isolated polynucleotides (I) encoding colon
CC tumour polypeptides (II). (I) is useful for stimulating an immune
CC response in a patient and treating colon cancer in a patient.
CC Oligonucleotides derived from (I) are useful for determining the presence
CC of cancer in a patient. (I) and (II) are useful in pharmaceutical
CC compositions, e.g. vaccines, and other compositions for the diagnosis
CC and treatment of colon cancer. A composition comprising a first component
CC selected from physiologically acceptable carriers and immunostimulants,
CC and an antigen-presenting cell expressing (II) is useful for inhibiting
CC development of cancer in a patient. (I) is useful in the design and
CC preparation of ribozyme molecules for inhibiting expression of tumour
CC polypeptides and (I). ABK54531-ABK55464 represent human colon cancer cDNA
CC sequences of the invention.

XX SQ Sequence 438 BP; 126 A; 118 C; 117 G; 77 T; 0 other;

Query Match 13.6%; Score 18; DB 24; Length 438;
Best Local Similarity 100.0%; Pred. No. 23;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 40 GATAAGGCCAAGCTGAAG 57

DB 99 GATAAGGCCAAGCTGAAG 116

RESULT 16

ABK55440

ID ABK55440 standard; cDNA; 439 BP.

XX AC ABK55440;

XX DT 18-JUN-2002 (first entry)

XX DE Human colon cancer-associated cDNA, SEQ ID No 910.

XX KW Human; colon cancer; immunogenic; vaccine; tumour; gene; ss.

XX OS Homo sapiens.

XX PN WO200212280-A2.

XX PD 14-FEB-2002.

XX PF 30-JUL-2001; 2001WO-US23826.

XX PR 03-AUG-2000; 2000US-223265P.

XX PR 02-OCT-2000; 2000US-237406P.

XX PR 20-MAR-2001; 2001US-277495P.

XX PR 03-JUL-2001; 2001US-302702P.

us-09-915-178-1.oli10.rng

Wed Jun 4 11:04:53 2003

CC tumour polypeptides (II). (I) is useful for stimulating an immune
 CC response in a patient and treating colon cancer in a patient.
 CC Oligonucleotides derived from (I) are useful for determining the presence
 CC of cancer in a patient. (I) and (II) are useful in pharmaceutical
 CC compositions, e.g. vaccines, and other compositions for the diagnosis
 CC and treatment of colon cancer. A composition comprising a first component
 CC selected from physiologically acceptable carriers and immunostimulants,
 CC and an antigen-presenting cell expressing (II) is useful for inhibiting
 CC development of cancer in a patient. (I) is useful in the design and
 CC preparation of ribozyme molecules for inhibiting expression of tumour
 CC polypeptides and (I). ABK54531-ABK55464 represent human colon cancer cDNA
 CC sequences of the invention.

XX SQ Sequence 443 BP; 128 A; 119 C; 119 G; 77 T; 0 other;
 Query Match 13.6%; Score 18; DB 24; Length 443;
 Best Local Similarity 100.0%; Pred. No. 23;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 40 GATAAGGCCAAGCTGAAG 57
 |||||
 DB 104 GATAAGGCCAAGCTGAAG 121

RESULT 18
 ABK55257
 ID ABK55257 standard; cDNA; 445 BP.
 AC ABK55257;
 XX
 XX 18-JUN-2002 (first entry)
 DT Human colon cancer-associated cDNA, SEQ ID No 727.
 DE Human; colon cancer; immunogenic; vaccine; tumour; gene; ss.
 DE Homo sapiens.
 KW WO200212280-A2.
 XX
 XX 14-FEB-2002.
 PD
 PD 30-JUL-2001; 2001WO-US23826.
 PF
 PF 03-AUG-2000; 2000US-223265P.
 PR 02-OCT-2000; 2000US-237406P.
 PR 20-MAR-2001; 2001US-277495P.
 PR 03-JUL-2001; 2001US-302702P.
 XX
 XX (CORI-) CORIXA CORP.
 PA
 XX Pyle RA, Xu J, Secretist H;
 PI
 XX WPI; 2002-257462/30.
 DR Novel polynucleotide encoding colon tumour polypeptides, useful as
 PT vaccines for treating colon cancers
 PT
 XX Claim 1; Page 355; 425pp; English.
 PS The invention relates to isolated polynucleotides (I) encoding colon
 XX tumour polypeptides (II). (I) is useful for stimulating an immune
 CC response in a patient and treating colon cancer in a patient.
 CC Oligonucleotides derived from (I) are useful for determining the presence
 CC of cancer in a patient. (I) and (II) are useful in pharmaceutical
 CC compositions, e.g. vaccines, and other compositions for the diagnosis
 CC and treatment of colon cancer. A composition comprising a first component
 CC selected from physiologically acceptable carriers and immunostimulants,
 CC and an antigen-presenting cell expressing (II) is useful for inhibiting
 CC development of cancer in a patient. (I) is useful in the design and
 CC preparation of ribozyme molecules for inhibiting expression of tumour
 CC polypeptides and (I). ABK54531-ABK55464 represent human colon cancer cDNA
 CC sequences of the invention.

PA (CORI-) CORIXA CORP.
 PI Pyle RA, Xu J, Secretist H;
 XX WPI; 2002-257462/30.
 DR Novel polynucleotide encoding colon tumour polypeptides, useful as
 PT vaccines for treating colon cancers
 PT
 XX Claim 1; Page 415; 425pp; English.
 PS The invention relates to isolated polynucleotides (I) encoding colon
 XX tumour polypeptides (II). (I) is useful for stimulating an immune
 CC response in a patient and treating colon cancer in a patient.
 CC Oligonucleotides derived from (I) are useful for determining the presence
 CC of cancer in a patient. (I) and (II) are useful in pharmaceutical
 CC compositions, e.g. vaccines, and other compositions for the diagnosis
 CC and treatment of colon cancer. A composition comprising a first component
 CC selected from physiologically acceptable carriers and immunostimulants,
 CC and an antigen-presenting cell expressing (II) is useful for inhibiting
 CC development of cancer in a patient. (I) is useful in the design and
 CC preparation of ribozyme molecules for inhibiting expression of tumour
 CC polypeptides and (I). ABK54531-ABK55464 represent human colon cancer cDNA
 CC sequences of the invention.

XX SQ Sequence 439 BP; 127 A; 116 C; 115 G; 78 T; 3 other;
 Query Match 13.6%; Score 18; DB 24; Length 439;
 Best Local Similarity 100.0%; Pred. No. 23;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 40 GATAAGGCCAAGCTGAAG 57
 |||||
 DB 94 GATAAGGCCAAGCTGAAG 111

RESULT 17
 ABK55189
 ID ABK55189 standard; cDNA; 443 BP.
 AC ABK55189;
 XX
 XX 18-JUN-2002 (first entry)
 DT Human colon cancer-associated cDNA, SEQ ID No 659.
 DE Human; colon cancer; immunogenic; vaccine; tumour; gene; ss.
 DE Homo sapiens.
 KW WO200212280-A2.
 XX
 XX 14-FEB-2002.
 PD
 PD 30-JUL-2001; 2001WO-US23826.
 PF
 PF 03-AUG-2000; 2000US-223265P.
 PR 02-OCT-2000; 2000US-237406P.
 PR 20-MAR-2001; 2001US-277495P.
 PR 03-JUL-2001; 2001US-302702P.
 XX
 XX (CORI-) CORIXA CORP.
 PA
 XX Pyle RA, Xu J, Secretist H;
 PI
 XX WPI; 2002-257462/30.
 DR Novel polynucleotide encoding colon tumour polypeptides, useful as
 PT vaccines for treating colon cancers
 PT
 XX Claim 1; Page 337; 425pp; English.
 PS The invention relates to isolated polynucleotides (I) encoding colon
 XX

```
XX SQ Sequence 445 BP; 130 A; 119 C; 119 G; 77 T; 0 other;
Query Match 13.6%; Score 18; DB 24; Length 445;
Best Local Similarity 100.0%; Pred. No. 23;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 40 GATAAGGCCAAGCTGAAG 57
|||||
Db 99 GATAAGGCCAAGCTGAAG 116

RESULT 19
ABN96661/c
ID ABN96661 standard; DNA; 446 BP.
XX AC ABN96661;
XX DT 13-AUG-2002 (first entry)
XX DE Gene #3159 used to diagnose liver cancer.
XX KW Gene; liver cancer; ds; hepatocellular carcinoma; hepatotropic;
XX KW metastatic liver tumour; cytostatic; expression profile; disease state;
XX KW disease progression; drug toxicity; drug efficacy; drug metabolism.
XX OS Homo sapiens.
XX PN WO200229103-A2.
XX PD 11-APR-2002.
XX PF 02-OCT-2001; 2001WO-US30589.
XX PR 02-OCT-2000; 2000US-237054P.
XX PA (GENE-) GENE LOGIC INC.
XX PI Horne D, Alvares C, Peres-Da-Silva S, Vockley JG;
XX DR WPI; 2002-426119/45.
XX DT Diagnosing and detecting the progression of liver cancer,
XX PT hepatocellular carcinoma or metastatic liver tumor in a patient,
XX PT involves detecting the level of expression of two or more genes in a
XX PT liver tissue sample
XX PS Claim 1; SEQ ID NO 3159; 298pp; English.
XX CC The invention relates to a novel method for diagnosing and detecting the
XX CC progression of liver cancer, hepatocellular carcinoma or metastatic liver
XX CC tumour in a patient, and differentiating metastatic liver cancer from
XX CC hepatocellular carcinoma in a patient, involving detecting the level of
XX CC expression of two or more genes represented in ABN93503-ABN97455 in a
XX CC tissue sample. The method of the invention has hepatotropic, and
XX CC cytosstatic activity. The method is useful for diagnosing and detecting
XX CC the progression of liver cancer, hepatocellular carcinoma and metastatic
XX CC liver carcinoma in a patient. The method is useful for identifying
XX CC expression profiles which serve as useful diagnostic markers as well as
XX CC markers that can be used to monitor disease states, disease progression,
XX CC drug toxicity, drug efficacy and drug metabolism.
XX CC Note: The sequence data for this patent did not form part of the printed
XX CC specification, but was obtained in electronic format directly from WIPO
XX CC at ftp.wipo.int/pub/published_pct_sequences.
XX SQ Sequence 446 BP; 74 A; 120 C; 122 G; 120 T; 10 other;

Query Match 13.6%; Score 18; DB 24; Length 446;
Best Local Similarity 100.0%; Pred. No. 23;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 40 GATAAGGCCAAGCTGAAG 57
|||||
```

```
Db 336 GATAAGGCCAAGCTGAAG 319

RESULT 20
ABK55421
ID ABK55421 standard; CDNA; 446 BP.
XX AC ABK55421;
XX DT 18-JUN-2002 (first entry)
XX DE Human colon cancer-associated CDNA, SEQ ID No 891.
XX DE Human; colon cancer; immunogenic; vaccine; tumour; gene; ss.
XX KW Homo sapiens.
XX OS WO200212280-A2.
XX PN 14-FEB-2002.
XX PD 30-JUL-2001; 2001WO-US23826.
XX PF 03-AUG-2000; 2000US-223265P.
XX PR 02-OCT-2000; 2000US-237406P.
XX PR 20-MAR-2001; 2001US-277495P.
XX PR 03-JUL-2001; 2001US-302702P.
XX PA (CORI-) CORIXA CORP.
XX PI Pyle RA, Xu J, Secrlist H;
XX DR WPI; 2002-257462/30.
XX DT Novel polynucleotide encoding colon tumour polypeptides, useful as
XX PT vaccines for treating colon cancers
XX PS Claim 1; Page 408; 425pp; English.
XX CC The invention relates to isolated polynucleotides (I) encoding colon
XX CC tumour polypeptides (II). (I) is useful for stimulating an immune
XX CC response in a patient and treating colon cancer in a patient.
XX CC Oligonucleotides derived from (I) are useful for determining the presence
XX CC of cancer in a patient. (I) and (II) are useful in pharmaceutical
XX CC compositions, e.g. vaccines, and other compositions for the diagnosis
XX CC and treatment of colon cancer. A composition comprising a first component
XX CC selected from physiologically acceptable carriers and immunostimulants,
XX CC and an antigen-presenting cell expressing (II) is useful for inhibiting
XX CC development of cancer in a patient. (I) is useful in the design and
XX CC preparation of ribozyme molecules for inhibiting expression of tumour
XX CC polypeptides and (I). ABK54531-ABK55464 represent human colon cancer CDNA
XX CC sequences of the invention.
XX SQ Sequence 446 BP; 141 A; 113 C; 111 G; 79 T; 2 other;

Query Match 13.6%; Score 18; DB 24; Length 446;
Best Local Similarity 100.0%; Pred. No. 23;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 40 GATAAGGCCAAGCTGAAG 57
|||||
Db 83 GATAAGGCCAAGCTGAAG 100

RESULT 21
ABK54652
ID ABK54652 standard; CDNA; 449 BP.
XX AC ABK54652;
XX DT 18-JUN-2002 (first entry)
XX DE Human colon cancer-associated CDNA, SEQ ID No 122.
```

us-09-915-178-1.olilo.rng

Wed Jun 4 11:04:53 2003

XX Human; colon cancer; immunogenic; vaccine; tumour; gene; ss.

XX Homo sapiens.

XX WO200212280-A2.

XX 14-FEB-2002.

XX 30-JUL-2001; 2001WO-US23826.

XX 03-AUG-2000; 2000US-223265P.

XX 02-OCT-2000; 2000US-237406P.

XX 20-MAR-2001; 2001US-277495P.

XX 03-JUL-2001; 2001US-302702P.

XX (CORI-) CORIXA CORP.

XX Pyle RA, Xu J, Secrlist H;

XX WPI; 2002-257462/30.

XX Novel polynucleotide encoding colon tumour polypeptides, useful as

XX vaccines for treating colon cancers

XX Claim 1; Page 185; 425pp; English.

XX The invention relates to isolated polynucleotides (I) encoding colon
XX tumour polypeptides (II). (I) is useful for stimulating an immune
XX response in a patient and treating colon cancer in a patient.
XX Oligonucleotides derived from (I) are useful for determining the presence
XX of cancer in a patient. (I) and (II) are useful in pharmaceutical
XX compositions, e.g. vaccines, and other compositions comprising a first component
XX and treatment of colon cancer. A composition comprising a first component
XX selected from physiologically acceptable carriers and immunostimulants,
XX and an antigen-presenting cell expressing (II) is useful for inhibiting
XX development of cancer in a patient. (I) is useful in the design and
XX preparation of ribozyme molecules for inhibiting expression of tumour
XX polypeptides and (I). ABK54531-ABK55464 represent human colon cancer cDNA
XX sequences of the invention.

XX Sequence 449 BP; 127 A; 120 C; 124 G; 78 T; 0 other;

XX Query Match 13.6%; Score 18; DB 24; Length 449;

XX Best Local Similarity 100.0%; Pred. No. 23;

XX Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

XX 40 GATAAGGCCCAAGCTGAAG 57

XX 112 GATAAGGCCCAAGCTGAAG 129

XX RESULT 22

XX ABK54720

XX ID ABK54720 standard; cDNA; 451 BP.

XX AC ABK54720;

XX 18-JUN-2002 (first entry)

XX Human colon cancer-associated cDNA, SEQ ID No 190.

XX Human; colon cancer; immunogenic; vaccine; tumour; gene; ss.

XX Homo sapiens.

XX WO200212280-A2.

XX 14-FEB-2002.

XX 30-JUL-2001; 2001WO-US23826.

XX 03-AUG-2000; 2000US-223265P.

PR 02-OCT-2000; 2000US-237406P.

PR 20-MAR-2001; 2001US-277495P.

PR 03-JUL-2001; 2001US-302702P.

XX (CORI-) CORIXA CORP.

XX Pyle RA, Xu J, Secrlist H;

XX WPI; 2002-257462/30.

XX Novel polynucleotide encoding colon tumour polypeptides, useful as
XX vaccines for treating colon cancers

XX Claim 1; Page 205; 425pp; English.

XX The invention relates to isolated polynucleotides (I) encoding colon
XX tumour polypeptides (II). (I) is useful for stimulating an immune
XX response in a patient and treating colon cancer in a patient.
XX Oligonucleotides derived from (I) are useful for determining the presence
XX of cancer in a patient. (I) and (II) are useful in pharmaceutical
XX compositions, e.g. vaccines, and other compositions comprising a first component
XX and treatment of colon cancer. A composition comprising a first component
XX selected from physiologically acceptable carriers and immunostimulants,
XX and an antigen-presenting cell expressing (II) is useful for inhibiting
XX development of cancer in a patient. (I) is useful in the design and
XX preparation of ribozyme molecules for inhibiting expression of tumour
XX polypeptides and (I). ABK54531-ABK55464 represent human colon cancer cDNA
XX sequences of the invention.

XX Sequence 451 BP; 127 A; 120 C; 126 G; 78 T; 0 other;

XX Query Match 13.6%; Score 18; DB 24; Length 451;

XX Best Local Similarity 100.0%; Pred. No. 23;

XX Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

XX 40 GATAAGGCCCAAGCTGAAG 57

XX 112 GATAAGGCCCAAGCTGAAG 129

XX RESULT 23

XX ABK55083

XX ID ABK55083 standard; cDNA; 451 BP.

XX AC ABK55083;

XX 18-JUN-2002 (first entry)

XX Human colon cancer-associated cDNA, SEQ ID No 553.

XX Human; colon cancer; immunogenic; vaccine; tumour; gene; ss.

XX Homo sapiens.

XX WO200212280-A2.

XX 14-FEB-2002.

XX 30-JUL-2001; 2001WO-US23826.

XX 03-AUG-2000; 2000US-223265P.

XX 02-OCT-2000; 2000US-237406P.

XX 20-MAR-2001; 2001US-277495P.

XX 03-JUL-2001; 2001US-302702P.

XX (CORI-) CORIXA CORP.

XX Pyle RA, Xu J, Secrlist H;

XX WPI; 2002-257462/30.

XX Novel polynucleotide encoding colon tumour polypeptides, useful as
XX vaccines for treating colon cancers

XX Claim 1; Page 308; 425pp; English.

XX The invention relates to isolated polynucleotides (I) encoding colon

CC tumour polypeptides (II). (I) is useful for stimulating an immune

CC response in a patient and treating colon cancer in a patient.

CC Oligonucleotides derived from (I) are useful for determining the presence

CC of cancer in a patient. (I) and (II) are useful in pharmaceutical

CC compositions, e.g. vaccines, and other compositions for the diagnosis

CC and treatment of colon cancer. A composition comprising a first component

CC selected from physiologically acceptable carriers and immunostimulants,

CC and an antigen-presenting cell expressing (II) is useful for inhibiting

CC development of cancer in a patient. (I) is useful in the design and

CC preparation of ribozyme molecules for inhibiting expression of tumour

CC polypeptides and (I). ABK54531-ABK55464 represent human colon cancer cDNA

CC sequences of the invention.

XX

SQ Sequence 451 BP; 129 A; 120 C; 124 G; 78 T; 0 other;

Query Match 13.6%; Score 18; DB 24; Length 451;

Best Local Similarity 100.0%; Pred. No. 23;

Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 40 GATAAGGCCCAAGCTGAAG 57

Db 112 GATAAGGCCCAAGCTGAAG 129

RESULT 24

ABK54878

ID ABK54878 standard; cDNA; 452 BP.

XX AC ABK54878;

XX

XX 18-JUN-2002 (first entry)

XX Human colon cancer-associated cDNA, SEQ ID No 348..

XX Human; colon cancer; immunogenic; vaccine; tumour; gene; ss.

XX Homo sapiens.

XX WO200212280-A2.

XX 14-FEB-2002.

XX 30-JUL-2001; 2001WO-US23826.

XX 03-AUG-2000; 2000US-223265P.

XX 02-OCT-2000; 2000US-237406P.

XX 20-MAR-2001; 2001US-277495P.

XX 03-JUL-2001; 2001US-302702P.

XX (CORI-) CORIXA CORP.

XX Pyle RA, Xu J, Secretist H;

XX WPI; 2002-257462/30.

XX Novel polynucleotide encoding colon tumour polypeptides, useful as

XX vaccines for treating colon cancers.

XX Claim 1; Page 249; 425pp; English.

XX The invention relates to isolated polynucleotides (I) encoding colon

CC tumour polypeptides (II). (I) is useful for stimulating an immune

CC response in a patient and treating colon cancer in a patient.

CC Oligonucleotides derived from (I) are useful for determining the presence

CC of cancer in a patient. (I) and (II) are useful in pharmaceutical

CC compositions, e.g. vaccines, and other compositions for the diagnosis

CC and treatment of colon cancer. A composition comprising a first component

CC selected from physiologically acceptable carriers and immunostimulants,

CC and an antigen-presenting cell expressing (II) is useful for inhibiting

CC development of cancer in a patient. (I) is useful in the design and

CC preparation of ribozyme molecules for inhibiting expression of tumour

CC polypeptides and (I). ABK54531-ABK55464 represent human colon cancer cDNA

CC sequences of the invention.

XX

SQ Sequence 452 BP; 129 A; 121 C; 124 G; 78 T; 0 other;

Query Match 13.6%; Score 18; DB 24; Length 452;

Best Local Similarity 100.0%; Pred. No. 23;

Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 40 GATAAGGCCCAAGCTGAAG 57

Db 113 GATAAGGCCCAAGCTGAAG 130

RESULT 25

AAC10525

ID AAC10525 standard; cDNA; 453 BP.

XX AC AAC10525;

XX

XX 06-OCT-2000 (first entry)

XX Human secreted protein 5' EST, SEQ ID NO: 14600.

XX Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;

XX gene therapy; chromosome mapping; ss.

XX Homo sapiens.

XX EP1033401-A2.

XX 06-SEP-2000.

XX 21-FEB-2000; 2000EP-0200610.

XX 26-FEB-1999; 99US-0122487.

XX (GEST) GENSET.

XX Dumas Milne Edwards J, Duclert A, Giordano J;

XX WPI; 2000-500381/45.

XX New nucleic acid that is a 5' expressed sequence tag (5' EST) for

XX obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for

XX diagnostic, forensic, gene therapy and chromosome mapping procedures -

XX Claim 1; SEQ ID 14600; 71pp + CD-ROM; English.

XX The present sequence is one of a large number of 5' ESTs derived from

XX mRNAs encoding secreted proteins. No ORF has yet been conclusively

XX identified within the present sequence. The 5' ESTs were prepared from

XX total human RNAs or polyA+ RNAs derived from 30 different tissues. EST

XX sequences usually correspond mainly to the 3' untranslated region (UTR)

XX of the mRNA because they are often obtained from oligo-dT primed cDNA

XX libraries. Such ESTs are not well suited for isolating cDNA sequences

XX derived from the 5' ends of mRNAs and even in those cases where longer

XX cDNA sequences have been obtained, the full 5' UTR is rarely included.

XX 5' ESTs are derived from mRNAs with intact 5' ends and can therefore be

XX used to obtain full length cDNAs and genomic DNAs. 5' ESTs are also used

XX in diagnostic, forensic, gene therapy and chromosome mapping procedures.

XX They are used to obtain upstream regulatory sequences and to design

XX expression and secretion vectors.

XX

SQ Sequence 453 BP; 111 A; 132 C; 130 G; 78 T; 2 other;

Query Match 13.6%; Score 18; DB 21; Length 453;

Best Local Similarity 100.0%; Pred. No. 23;

Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 40 GATAAGGCCCAAGCTGAAG 57

us-09-915-178-1.0110.rng

Wed Jun 4 11:04:53 2003

```

XX DT 18-JUN-2002 (first entry)
XX DE Human colon cancer-associated cDNA, SEQ ID NO 274.
XX KW Human; colon cancer; immunogenic; vaccine; tumour; gene; ss.
XX OS Homo sapiens.
XX PN WO200212280-A2.
XX PD 14-FEB-2002.
XX PF 30-JUL-2001; 2001WO-US23826.
XX PR 03-AUG-2000; 2000US-223265P.
XX PR 02-OCT-2000; 2000US-237406P.
XX PR 20-MAR-2001; 2001US-277495P.
XX PR 03-JUL-2001; 2001US-302702P.
XX PA (CORI-) CORIXA CORP.
XX PI Pyle RA, Xu J, Secretist H;
XX DR WPI; 2002-257462/30.
XX PT Novel polynucleotide encoding colon tumour polypeptides, useful as
XX PT vaccines for treating colon cancers
XX PS Claim 1; Page 228; 425pp; English.
XX CC The invention relates to isolated polynucleotides (I) encoding colon
XX CC tumour polypeptides (II). (I) is useful for stimulating an immune
XX CC response in a patient and treating colon cancer in a patient.
XX CC Oligonucleotides derived from (I) are useful for determining the presence
XX CC of cancer in a patient. (I) and (II) are useful in pharmaceutical
XX CC compositions, e.g. vaccines, and other compositions for the diagnosis
XX CC and treatment of colon cancer. A composition comprising a first component
XX CC selected from physiologically acceptable carriers and immunostimulants,
XX CC and an antigen-presenting cell expressing (II) is useful for inhibiting
XX CC development of cancer in a patient. (I) is useful in the design and
XX CC preparation of ribozyme molecules for inhibiting expression of tumour
XX CC polypeptides and (I). ABK54531-ABK55464 represent human colon cancer cDNA
XX CC sequences of the invention.
XX SQ Sequence 453 BP; 130 A; 121 C; 124 G; 78 T; 0 other;
XX Query Match 13.6%; Score 18; DB 24; Length 453;
XX Best Local Similarity 100.0%; Pred. No. 23;
XX Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 40 GATAAGGCCCAAGCTGAAG 57
DB 114 GATAAGGCCCAAGCTGAAG 131
RESULT 28
ABK54803
ID ABK54803 standard; cDNA; 455 BP.
XX AC ABK54803;
XX DT 18-JUN-2002 (first entry)
XX DE Human colon cancer-associated cDNA, SEQ ID NO 273.
XX KW Human; colon cancer; immunogenic; vaccine; tumour; gene; ss.
XX OS Homo sapiens.
XX PN WO200212280-A2.
XX PD 14-FEB-2002.
XX AC ABK54804;

1147 GATAAGGCCCAAGCTGAAG 164
RESULT 26
ABN96522
ID ABN96522 standard; DNA; 453 BP.
XX AC ABN96522;
XX DT 13-AUG-2002 (first entry)
XX DE Gene #3020 used to diagnose liver cancer.
XX KW Gene; liver cancer; ds; hepatocellular carcinoma; hepatotropic;
XX KW metastatic liver tumour; cytostatic; expression profile; disease state;
XX KW disease progression; drug toxicity; drug efficacy; drug metabolism.
XX OS Homo sapiens.
XX PN WO200229103-A2.
XX PD 11-APR-2002.
XX PF 02-OCT-2001; 2001WO-US30589.
XX PR 02-OCT-2000; 2000US-237054P.
XX PA (GENE-) GENE LOGIC INC.
XX PI Horne D, Alvares C, Peres-Da-Silva S, Vockley JG;
XX DR WPI; 2002-426119/45.
XX PT Diagnosing and detecting the progression of liver cancer,
XX PT hepatocellular carcinoma or metastatic liver tumor in a patient,
XX PT involves detecting the level of expression of two or more genes in a
XX PT liver tissue sample
XX PS Claim 1; SEQ ID NO 3020; 298pp; English.
XX CC The invention relates to a novel method for diagnosing and detecting the
XX CC progression of liver cancer, hepatocellular carcinoma or metastatic liver
XX CC tumour in a patient, and differentiating metastatic liver cancer from
XX CC hepatocellular carcinoma in a patient, involving detecting the level of
XX CC expression of two or more genes represented in ABN93503-ABN97455 in a
XX CC tissue sample. The method of the invention has hepatotropic, and
XX CC cytostatic activity. The method is useful for diagnosing and detecting
XX CC the progression of liver cancer, hepatocellular carcinoma and metastatic
XX CC liver carcinoma in a patient. The method is useful for identifying
XX CC expression profiles which serve as useful diagnostic markers as well as
XX CC markers that can be used to monitor disease states, disease progression,
XX CC drug toxicity, drug efficacy and drug metabolism.
XX CC Note: The sequence data for this patent did not form part of the printed
XX CC specification, but was obtained in electronic format directly from WIPO
XX CC at ftp.wipo.int/pub/published_pct_sequences.
XX SQ Sequence 453 BP; 132 A; 119 C; 124 G; 78 T; 0 other;
XX Query Match 13.6%; Score 18; DB 24; Length 453;
XX Best Local Similarity 100.0%; Pred. No. 23;
XX Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 40 GATAAGGCCCAAGCTGAAG 57
DB 105 GATAAGGCCCAAGCTGAAG 122
RESULT 27
ABK54804
ID ABK54804 standard; cDNA; 453 BP.
XX AC ABK54804;
```

XX PF 30-JUL-2001; 2001WO-US23826.
XX PR 03-AUG-2000; 2000US-223265P.
XX PR 02-OCT-2000; 2000US-237406P.
XX PR 20-MAR-2001; 2001US-277495P.
XX PR 03-JUL-2001; 2001US-302702P.
XX PA (CORI-) CORIXA CORP.
XX PI Pyle RA, Xu J, Secrlist H;
XX DR WPI; 2002-257462/30.
XX PT Novel polynucleotide encoding colon tumour polypeptides, useful as
XX PT vaccines for treating colon cancers
XX PS Claim 1; Page 228; 425pp; English.
XX CC The invention relates to isolated polynucleotides (I) encoding colon
XX CC tumour polypeptides (II). (I) is useful for stimulating an immune
XX CC response in a patient and treating colon cancer in a patient.
XX CC Oligonucleotides derived from (I) are useful for determining the presence
XX CC of cancer in a patient. (I) and (II) are useful for determining the presence
XX CC of cancer in a patient. (I) and (II) are useful in pharmaceutical
XX CC compositions, e.g. vaccines, and other compositions comprising a first component
XX CC selected from physiologically acceptable carriers and immunostimulants,
XX CC and an antigen-presenting cell expressing (II) is useful for inhibiting
XX CC development of cancer in a patient. (I) is useful in the design and
XX CC preparation of ribozyme molecules for inhibiting expression of tumour
XX CC polypeptides and (I). ABK54531-ABK5464 represent human colon cancer cdna
XX CC sequences of the invention.
XX SQ Sequence 455 BP; 130 A; 121 C; 126 G; 78 T; 0 other;
XX
XX Query Match 13.6%; Score 18; DB 24; Length 455;
XX Best Local Similarity 100.0%; Pred. No. 23;
XX Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 40 GATAGGCCCAAGCTGAAG 57
XX ||||||||||||||||
XX DB 112 GATAGGCCCAAGCTGAAG 129
XX
XX RESULT 29
XX ABK54749
XX ID ABK54749 standard; cdna; 456 BP.
XX AC ABK54749;
XX DT 18-JUN-2002 (first entry)
XX DE Human colon cancer-associated cdna, SEQ ID No 219.
XX KW Human; colon cancer; immunogenic; vaccine; tumour; gene; ss.
XX OS Homo sapiens.
XX PN WO200212280-A2.
XX PD 14-FEB-2002.
XX PF 30-JUL-2001; 2001WO-US23826.
XX PR 03-AUG-2000; 2000US-223265P.
XX PR 02-OCT-2000; 2000US-237406P.
XX PR 20-MAR-2001; 2001US-277495P.
XX PR 03-JUL-2001; 2001US-302702P.
XX PA (CORI-) CORIXA CORP.
XX PI Pyle RA, Xu J, Secrlist H;
XX PX

DR WPI; 2002-257462/30.
XX Novel polynucleotide encoding colon tumour polypeptides, useful as
XX PT vaccines for treating colon cancers
XX PS Claim 1; Page 212-213; 425pp; English.
XX CC The invention relates to isolated polynucleotides (I) encoding colon
XX CC tumour polypeptides (II). (I) is useful for stimulating an immune
XX CC response in a patient and treating colon cancer in a patient.
XX CC Oligonucleotides derived from (I) are useful for determining the presence
XX CC of cancer in a patient. (I) and (II) are useful in pharmaceutical
XX CC compositions, e.g. vaccines, and other compositions comprising a first component
XX CC and treatment of colon cancer. A composition comprising a first component
XX CC selected from physiologically acceptable carriers and immunostimulants,
XX CC and an antigen-presenting cell expressing (II) is useful for inhibiting
XX CC development of cancer in a patient. (I) is useful in the design and
XX CC preparation of ribozyme molecules for inhibiting expression of tumour
XX CC polypeptides and (I). ABK54531-ABK5464 represent human colon cancer cdna
XX CC sequences of the invention.
XX SQ Sequence 456 BP; 131 A; 122 C; 125 G; 78 T; 0 other;
XX
XX Query Match 13.6%; Score 18; DB 24; Length 456;
XX Best Local Similarity 100.0%; Pred. No. 23;
XX Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 40 GATAGGCCCAAGCTGAAG 57
XX ||||||||||||||||
XX DB 117 GATAGGCCCAAGCTGAAG 134
XX
XX RESULT 30
XX ABK54586
XX ID ABK54586 standard; cdna; 465 BP.
XX AC ABK54586;
XX DT 18-JUN-2002 (first entry)
XX DE Human colon cancer-associated cdna, SEQ ID No 56.
XX KW Human; colon cancer; immunogenic; vaccine; tumour; gene; ss.
XX OS Homo sapiens.
XX PN WO200212280-A2.
XX PD 14-FEB-2002.
XX PF 30-JUL-2001; 2001WO-US23826.
XX PR 03-AUG-2000; 2000US-223265P.
XX PR 02-OCT-2000; 2000US-237406P.
XX PR 20-MAR-2001; 2001US-277495P.
XX PR 03-JUL-2001; 2001US-302702P.
XX PA (CORI-) CORIXA CORP.
XX PI Pyle RA, Xu J, Secrlist H;
XX PX WPI; 2002-257462/30.
XX Novel polynucleotide encoding colon tumour polypeptides, useful as
XX PT vaccines for treating colon cancers
XX PS Claim 1; Page 166; 425pp; English.
XX CC The invention relates to isolated polynucleotides (I) encoding colon
XX CC tumour polypeptides (II). (I) is useful for stimulating an immune
XX CC response in a patient and treating colon cancer in a patient.
XX CC Oligonucleotides derived from (I) are useful for determining the presence
XX CC of cancer in a patient. (I) and (II) are useful in pharmaceutical

compositions, e.g. vaccines, and other compositions for the diagnosis and treatment of colon cancer. A composition comprising a first component selected from physiologically acceptable carriers and immunostimulants, and an antigen-presenting cell expressing (ii) is useful for inhibiting development of cancer in a patient. (i) is useful in the design and preparation of ribozyme molecules for inhibiting expression of tumour polypeptides and (i). ABK54531-ABK5464 represent human colon cancer cDNA sequences of the invention.

Sequence 465 BP; 130 A; 124 C; 124 G; 84 T; 3 other;

Query Match 13.6%; Score 18; DB 24; Length 465;
Best Local Similarity 100.0%; Pred. No. 23;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 40 GATAAGGCCCAAGCTGAAG 57
|||||

DB 101 GATAAGGCCCAAGCTGAAG 118
|||||

RESULT 31
AAC10524
ID AAC10524 standard; cDNA; 488 BP.
XX
AC AAC10524;
XX
XX 06-OCT-2000 (first entry)
XX
XX Human secreted protein 5' EST, SEQ ID NO: 14599.

XX Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
KW gene therapy; chromosome mapping; ss.
XX Homo sapiens.
XX
XX EP1033401-A2.
XX
XX 06-SEP-2000.

XX 21-FEB-2000; 2000EP-0200610.
XX
XX 26-FEB-1999; 99US-0122487.
XX
XX (GIST) GENSET.
XX
XX Dumas Milne Edwards J, Duclert A, Giordano J;
XX
XX WPI; 2000-500381/45.

New nucleic acid that is a 5' expressed sequence tag (5' EST) for obtaining cDNAs and genomic DNAs that correspond to 5' ESTs and for diagnostic, forensic, gene therapy and chromosome mapping procedures -
Claim 1; SEQ ID 14599; 71pp + CD-ROM; English.

The present sequence is one of a large number of 5' ESTs derived from mRNAs encoding secreted proteins. No ORF has yet been conclusively identified within the present sequence. The 5' ESTs were prepared from total human RNAs or polyA+ RNAs derived from 30 different tissues. EST sequences usually correspond mainly to the 3' untranslated region (UTR) of the mRNA because they are often obtained from oligo-dT primed cDNA libraries. Such ESTs are not well suited for isolating cDNA sequences derived from the 5' ends of mRNAs and even in those cases where longer cDNA sequences have been obtained, the full 5' UTR is rarely included. 5' ESTs are derived from mRNAs with intact 5' ends and can therefore be used to obtain full length cDNAs and genomic DNAs. 5' ESTs are also used in diagnostic, forensic, gene therapy and chromosome mapping procedures. They are used to obtain upstream regulatory sequences and to design expression and secretion vectors.

Sequence 488 BP; 135 A; 127 C; 136 G; 90 T; 0 other;
Query Match 13.6%; Score 18; DB 21; Length 488;

Best Local Similarity 100.0%; Pred. No. 23;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 40 GATAAGGCCCAAGCTGAAG 57
|||||

DB 147 GATAAGGCCCAAGCTGAAG 164
|||||

RESULT 32
AAD03640
ID AAD03640 standard; DNA; 491 BP.
XX
AC AAD03640;
XX
XX 19-JUN-2001 (first entry)
XX
XX Human thymosin b-10 gene from clone NL_33.

XX Human; natural antisense mRNA enrichment; antisense-based therapy;
KW thymosin b-10; clone NL_33; ds.
XX Homo sapiens.
XX
XX WO200125488-A2.
XX
XX 12-APR-2001.

XX 06-OCT-2000; 2000WO-US27557.
XX
XX 06-OCT-1999; 99US-0157843.
XX
XX (QUAR-) QUARK BIOTECH INC.
XX
XX Gilad S, Einat P, Grossman A;
XX
XX WPI; 2001-266326/27.

Enrichment and detection of natural antisense mRNA comprises generating double stranded hybrid cDNA using a polymerase with an exonuclease activity, amplifying using a DT primer and cloning -
Disclosure; Page 30; 37pp; English.

The invention relates to a method for enrichment of natural antisense messenger RNA. This method involves generating a population of cDNA from mRNA, incubating the generated cDNA to produce double stranded hybrid DNA molecules consisting of sense and antisense cDNA, treating the hybrid molecules using DNA polymerase with an exonuclease activity, amplifying the double stranded molecule using a deoxythymidine (dT) primer and cloning the amplified double stranded molecule. This method is useful for enrichment of natural antisense mRNA from any natural source of RNA. It is used to detect whether mRNAs have a natural anti-sense counterpart. The method provides a basis for finding new genes with important cellular regulatory roles or new regulatory information for known genes and provides a starting material for development of an antisense-based therapeutic to treat a disease in which down regulation or inhibition of the sense gene or transcript is involved.

The present sequence is human thymosin b-10 gene from clone NL_33. This clone is obtained from the antisense enriched library and is individually confirmed for the presence of matching antisense mRNA.

Sequence 491 BP; 150 A; 120 C; 123 G; 98 T; 0 other;
Query Match 13.6%; Score 18; DB 22; Length 491;
Best Local Similarity 100.0%; Pred. No. 23;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 40 GATAAGGCCCAAGCTGAAG 57
|||||


```
PN WO200160860-A2.
XX 23-AUG-2001.
XX 20-FEB-2001; 2001WO-US05171.
XX 17-FEB-2000; 2000US-183319P.
XX 16-MAR-2000; 2000US-189862P.
XX 25-MAY-2000; 2000US-207454P.
XX 09-JUN-2000; 2000US-211314P.
XX 18-JUL-2000; 2000US-219007P.
XX 13-DEC-2000; 2000US-255281P.
XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX Schlegel R, Endege WO, Monahan JE;
XX WPI; 2001-662795/76.
XX Novel isolated nucleic acid molecule associated with cancerous state of
XX prostate cells and correlating with presence of prostate cancer, useful
XX for detecting presence of prostate cancer, stage of prostate cancer
XX Claim 1; Page 8549; 11750pp; English.
XX The invention relates to an isolated nucleic acid molecule (I) comprising
XX a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
XX specification or its complement. (I) is useful for:
XX (a) assessing whether a patient is afflicted with prostate cancer;
XX (b) monitoring the progression of prostate cancer in a patient;
XX (c) assessing the efficacy of a test compound to inhibit prostate
XX cancer in a patient;
XX (d) assessing the efficacy of a therapy for inhibiting prostate cancer
XX in a patient;
XX (e) selecting a composition for inhibiting prostate cancer in a patient;
XX (f) assessing the prostate cell carcinogenic potential of a compound;
XX (g) determining whether prostate cancer has metastasized in a patient;
XX (h) assessing the aggressiveness or indolence of prostate cancer in a
XX patient;
XX (I) is also useful as a pharmacodynamic or pharmacogenomic marker.
XX Sequence 534 BP; 137 A; 131 C; 113 G; 153 T; 0 other;
XX Query Match 13.6%; Score 18; DB 23; Length 534;
XX Best Local Similarity 100.0%; Pred. No. 23;
XX Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 40 GATAAGGCCCAAGCTGAAG 57
DB 377 GATAAGGCCCAAGCTGAAG 394
RESULT 36
ABK54689
ID ABK54689 standard; cDNA; 559 BP.
XX AC ABK54689;
XX DT 18-JUN-2002 (first entry)
XX Human colon cancer-associated cDNA, SEQ ID NO 159.
XX Human; colon cancer; immunogenic; vaccine; tumour; gene; ss.
XX Homo sapiens.
XX WO200212280-A2.
XX 14-FEB-2002.
XX 30-JUL-2001; 2001WO-US23826.
XX 03-AUG-2000; 2000US-223265P.
XX 02-OCT-2000; 2000US-237406P.
XX 20-MAR-2001; 2001US-277495P.
XX 03-JUL-2001; 2001US-302702P.
XX (CORI-) CORIXA CORP.
XX Pyle RA, Xu J, Secrlist H;
XX WPI; 2002-257462/30.
XX Novel polynucleotide encoding colon tumour polypeptides, useful as
XX vaccines for treating colon cancers
XX Claim 1; Page 196; 425pp; English.
XX The invention relates to isolated polynucleotides (I) encoding colon
XX tumour polypeptides (II). (I) is useful for stimulating an immune
XX response in a patient and treating colon cancer in a patient.
XX Oligonucleotides derived from (I) are useful for determining the presence
XX of cancer in a patient. (I) and (II) are useful in pharmaceutical
XX compositions, e.g. vaccines, and other compositions for the diagnosis
XX and treatment of colon cancer. A composition comprising a first component
XX selected from physiologically acceptable carriers and immunostimulants,
XX and an antigen-presenting cell expressing (II) is useful for inhibiting
XX development of cancer in a patient. (I) is useful in the design and
XX preparation of ribozyme molecules for inhibiting expression of tumour
XX polypeptides and (I). ABK54531-ABK55464 represent human colon cancer cDNA
XX sequences of the invention.
XX Sequence 559 BP; 118 A; 196 C; 144 G; 99 T; 2 other;
XX Query Match 13.6%; Score 18; DB 24; Length 559;
XX Best Local Similarity 100.0%; Pred. No. 23;
XX Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 40 GATAAGGCCCAAGCTGAAG 57
DB 90 GATAAGGCCCAAGCTGAAG 107
RESULT 37
ABK55328
ID ABK55328 standard; cDNA; 693 BP.
XX AC ABK55328;
XX DT 18-JUN-2002 (first entry)
XX Human colon cancer-associated cDNA, SEQ ID NO 798.
XX Human; colon cancer; immunogenic; vaccine; tumour; gene; ss.
XX Homo sapiens.
XX WO200212280-A2.
XX 14-FEB-2002.
XX 30-JUL-2001; 2001WO-US23826.
XX 03-AUG-2000; 2000US-223265P.
XX 02-OCT-2000; 2000US-237406P.
XX 20-MAR-2001; 2001US-277495P.
XX 03-JUL-2001; 2001US-302702P.
XX (CORI-) CORIXA CORP.
XX Pyle RA, Xu J, Secrlist H;
XX WPI; 2002-257462/30.
XX Novel polynucleotide encoding colon tumour polypeptides, useful as
XX vaccines for treating colon cancers
XX Claim 1; Page 196; 425pp; English.
XX The invention relates to isolated polynucleotides (I) encoding colon
XX tumour polypeptides (II). (I) is useful for stimulating an immune
XX response in a patient and treating colon cancer in a patient.
XX Oligonucleotides derived from (I) are useful for determining the presence
XX of cancer in a patient. (I) and (II) are useful in pharmaceutical
XX compositions, e.g. vaccines, and other compositions for the diagnosis
XX and treatment of colon cancer. A composition comprising a first component
XX selected from physiologically acceptable carriers and immunostimulants,
XX and an antigen-presenting cell expressing (II) is useful for inhibiting
XX development of cancer in a patient. (I) is useful in the design and
XX preparation of ribozyme molecules for inhibiting expression of tumour
XX polypeptides and (I). ABK54531-ABK55464 represent human colon cancer cDNA
XX sequences of the invention.
XX Sequence 559 BP; 118 A; 196 C; 144 G; 99 T; 2 other;
XX Query Match 13.6%; Score 18; DB 24; Length 559;
XX Best Local Similarity 100.0%; Pred. No. 23;
XX Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 40 GATAAGGCCCAAGCTGAAG 57
DB 90 GATAAGGCCCAAGCTGAAG 107
```

XX Claim 1; Page 376; 425pp; English.

XX The invention relates to isolated polynucleotides (I) encoding colon

CC tumour polypeptides (II). (I) is useful for stimulating an immune

CC response in a patient and treating colon cancer in a patient.

CC Oligonucleotides derived from (I) are useful for determining the presence

CC of cancer in a patient. (I) and (II) are useful in pharmaceutical

CC compositions, e.g. vaccines, and other compositions for the diagnosis

CC and treatment of colon cancer. A composition comprising a first component

CC selected from physiologically acceptable carriers and immunostimulants,

CC and an antigen-presenting cell expressing (II) is useful for inhibiting

CC development of cancer in a patient. (I) is useful in the design and

CC preparation of ribozyme molecules for inhibiting expression of tumour

CC polypeptides and (I). ABK54531-ABK5464 represent human colon cancer cDNA

CC sequences of the invention.

XX SQ Sequence 693 BP; 203 A; 168 C; 172 G; 130 T; 20 other;

Query Match 13.6%; Score 18; DB 24; Length 693;

Best Local Similarity 100.0%; Pred. No. 23;

Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 40 GATAAGGCCAAGCTGAAG 57

DB 80 GATAAGGCCAAGCTGAAG 97

|||||

RESULT 38

ABV25565

ID ABV25565 standard; cDNA; 736 BP.

XX AC ABV25565;

XX DT 16-SEP-2002 (first entry)

XX DE Human prostate expression marker cDNA 25556.

XX KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;

XX KW pharmacogenomic marker; gene; ss.

XX OS Homo sapiens.

XX FN WO200160860-A2.

XX PD 23-AUG-2001.

XX PF 20-FEB-2001; 2001WO-US05171.

XX PR 17-FEB-2000; 2000US-183319P.

XX PR 16-MAR-2000; 2000US-189862P.

XX PR 25-MAY-2000; 2000US-207454P.

XX PR 09-JUN-2000; 2000US-211314P.

XX PR 18-JUL-2000; 2000US-219007P.

XX PR 13-DEC-2000; 2000US-255281P.

XX PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

XX PI Schlegel R, Endege WO, Monahan JE;

XX DR WPI; 2001-662795/76.

XX PT Novel isolated nucleic acid molecule associated with cancerous state of

XX PT prostate cells and correlating with presence of prostate cancer, useful

XX PT for detecting presence of prostate cancer, stage of prostate cancer -

XX PS Claim 1; Page 5098; 11750pp; English.

XX The invention relates to an isolated nucleic acid molecule (I) comprising

CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the

CC specification or its complement. (I) is useful for:

CC (a) assessing whether a patient is afflicted with prostate cancer;

CC (b) monitoring the progression of prostate cancer in a patient;

XX PS Claim 1; Page 5098; 11750pp; English.

XX The invention relates to an isolated nucleic acid molecule (I) comprising

CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the

CC specification or its complement. (I) is useful for:

CC (a) assessing whether a patient is afflicted with prostate cancer;

CC (b) monitoring the progression of prostate cancer in a patient;

CC (c) assessing the efficacy of a test compound to inhibit prostate

CC cancer in a patient;

CC (d) assessing the efficacy of a therapy for inhibiting prostate cancer

CC in a patient;

CC (e) selecting a composition for inhibiting prostate cancer in a patient;

CC (f) assessing the prostate cell carcinogenic potential of a compound;

CC (g) determining whether prostate cancer has metastasized in a patient;

CC (h) assessing the aggressiveness or indolence of prostate cancer in a

CC patient;

CC (I) is also useful as a pharmacodynamic or pharmacogenomic marker.

XX SQ Sequence 736 BP; 165 A; 245 C; 188 G; 138 T; 0 other;

Query Match 13.6%; Score 18; DB 23; Length 736;

Best Local Similarity 100.0%; Pred. No. 23;

Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 40 GATAAGGCCAAGCTGAAG 57

DB 127 GATAAGGCCAAGCTGAAG 144

|||||

RESULT 39

ABV03577

ID ABV03577 standard; cDNA; 792 BP.

XX AC ABV03577;

XX DT 13-SEP-2002 (first entry)

XX DE Human prostate expression marker cDNA 3568.

XX KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;

XX KW pharmacogenomic marker; gene; ss.

XX OS Homo sapiens.

XX FN WO200160860-A2.

XX PD 23-AUG-2001.

XX PF 20-FEB-2001; 2001WO-US05171.

XX PR 17-FEB-2000; 2000US-183319P.

XX PR 16-MAR-2000; 2000US-189862P.

XX PR 25-MAY-2000; 2000US-207454P.

XX PR 09-JUN-2000; 2000US-211314P.

XX PR 18-JUL-2000; 2000US-219007P.

XX PR 13-DEC-2000; 2000US-255281P.

XX PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

XX PI Schlegel R, Endege WO, Monahan JE;

XX DR WPI; 2001-662795/76.

XX PT Novel isolated nucleic acid molecule associated with cancerous state of

XX PT prostate cells and correlating with presence of prostate cancer, useful

XX PT for detecting presence of prostate cancer, stage of prostate cancer -

XX PS Claim 1; Page 645; 11750pp; English.

XX The invention relates to an isolated nucleic acid molecule (I) comprising

CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the

CC specification or its complement. (I) is useful for:

CC (a) assessing whether a patient is afflicted with prostate cancer;

CC (b) monitoring the progression of prostate cancer in a patient;

CC (c) assessing the efficacy of a test compound to inhibit prostate

CC cancer in a patient;

CC (d) assessing the efficacy of a therapy for inhibiting prostate cancer

CC in a patient;

CC (e) selecting a composition for inhibiting prostate cancer in a patient;

CC (f) assessing the prostate cell carcinogenic potential of a compound;

us-09-915-178-1.olilo.rng

Wed Jun 4 11:04:53 2003

CC (g) determining whether prostate cancer has metastasized in a patient;
 CC (h) assessing the aggressiveness or indolence of prostate cancer in a
 CC patient;
 CC (I) is also useful as a pharmacodynamic or pharmacogenomic marker.
 XX Sequence 792 BP; 207 A; 187 C; 154 G; 190 T; 54 other;
 SQ Query Match 13.6%; Score 18; DB 23; Length 792;
 Best Local Similarity 100.0%; Pred. No. 23;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 40 GATAGGCCCAAGCTGAAG 57
 DB 381 GATAGGCCCAAGCTGAAG 398
 RESULT 40
 ABV12746
 ID ABV12746 standard; cDNA; 836 BP.
 AC ABV12746;
 XX 13-SEP-2002 (first entry)
 DT Human prostate expression marker cDNA 12737.
 DE Human prostate expression marker cDNA 12737.
 XX Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
 KW pharmacogenomic marker; gene; ss.
 XX Homo sapiens.
 OS WO200160860-A2.
 XX 23-AUG-2001.
 PD 20-FEB-2001; 2001WO-US05171.
 PF 17-FEB-2000; 2000US-183319P.
 XX 16-MAR-2000; 2000US-189862P.
 PR 25-MAY-2000; 2000US-207454P.
 PR 09-JUN-2000; 2000US-211314P.
 PR 18-JUL-2000; 2000US-219007P.
 PR 13-DEC-2000; 2000US-255281P.
 XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
 PA Schlegel R, Endege WO, Monahan JE;
 PI WPI; 2001-662795/76.
 XX Novel isolated nucleic acid molecule associated with cancerous state of
 XX prostate cells and correlating with presence of prostate cancer, useful
 XX for detecting presence of prostate cancer, stage of prostate cancer -
 XX Claim 1; Page 2099; 11750pp; English.
 PS The invention relates to an isolated nucleic acid molecule (I) comprising
 XX a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
 CC specification or its complement. (I) is useful for:
 CC (a) assessing whether a patient is afflicted with prostate cancer;
 CC (b) monitoring the progression of prostate cancer in a patient;
 CC (c) assessing the efficacy of a test compound to inhibit prostate
 CC cancer in a patient;
 CC (d) assessing the efficacy of a therapy for inhibiting prostate cancer
 CC in a patient;
 CC (e) selecting a composition for inhibiting prostate cancer in a patient;
 CC (f) assessing the prostate cell carcinogenic potential of a compound;
 CC (g) determining whether prostate cancer has metastasized in a patient;
 CC (h) assessing the aggressiveness or indolence of prostate cancer in a
 CC patient;
 CC (I) is also useful as a pharmacodynamic or pharmacogenomic marker.
 XX Sequence 836 BP; 226 A; 201 C; 178 G; 221 T; 10 other;
 SQ

Query Match 13.6%; Score 18; DB 23; Length 836;
 Best Local Similarity 100.0%; Pred. No. 23;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 40 GATAGGCCCAAGCTGAAG 57
 DB 402 GATAGGCCCAAGCTGAAG 419
 RESULT 41
 ABV21155
 ID ABV21155 standard; cDNA; 1201 BP.
 AC ABV21155;
 XX 13-SEP-2002 (first entry)
 DT Human prostate expression marker cDNA 21146.
 DE Human prostate expression marker cDNA 21146.
 XX Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
 KW pharmacogenomic marker; gene; ss.
 XX Homo sapiens.
 OS WO200160860-A2.
 XX 23-AUG-2001.
 PD 20-FEB-2001; 2001WO-US05171.
 PF 17-FEB-2000; 2000US-183319P.
 XX 16-MAR-2000; 2000US-189862P.
 PR 25-MAY-2000; 2000US-207454P.
 PR 09-JUN-2000; 2000US-211314P.
 PR 18-JUL-2000; 2000US-219007P.
 PR 13-DEC-2000; 2000US-255281P.
 XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
 PA Schlegel R, Endege WO, Monahan JE;
 PI WPI; 2001-662795/76.
 XX Novel isolated nucleic acid molecule associated with cancerous state of
 XX prostate cells and correlating with presence of prostate cancer, useful
 XX for detecting presence of prostate cancer, stage of prostate cancer -
 XX Claim 1; Page 3506; 11750pp; English.
 PS The invention relates to an isolated nucleic acid molecule (I) comprising
 XX a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
 CC specification or its complement. (I) is useful for:
 CC (a) assessing whether a patient is afflicted with prostate cancer;
 CC (b) monitoring the progression of prostate cancer in a patient;
 CC (c) assessing the efficacy of a test compound to inhibit prostate
 CC cancer in a patient;
 CC (d) assessing the efficacy of a therapy for inhibiting prostate cancer
 CC in a patient;
 CC (e) selecting a composition for inhibiting prostate cancer in a patient;
 CC (f) assessing the prostate cell carcinogenic potential of a compound;
 CC (g) determining whether prostate cancer has metastasized in a patient;
 CC (h) assessing the aggressiveness or indolence of prostate cancer in a
 CC patient;
 CC (I) is also useful as a pharmacodynamic or pharmacogenomic marker.
 XX Sequence 1201 BP; 355 A; 291 C; 289 G; 265 T; 1 other;
 SQ Query Match 13.6%; Score 18; DB 23; Length 1201;
 Best Local Similarity 100.0%; Pred. No. 23;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 40 GATAGGCCCAAGCTGAAG 57

|||||
401 GATAAGGCCAAGCTGAAG 418

RESULT 42

ABV22339

ID ABV22339 standard; cDNA; 1201 BP.

XX

AC ABV22339;

XX

DT 13-SEP-2002 (first entry)

XX

DE Human prostate expression marker cDNA 22330.

XX

KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
pharmacogenomic marker; gene; ss.

XX

OS Homo sapiens.

XX

FN WO200160860-A2.

XX

PD 23-AUG-2001.

XX

PF 20-FEB-2001; 2001WO-US05171.

XX

PR 17-FEB-2000; 2000US-183319P.

XX

PR 16-MAR-2000; 2000US-189862P.

XX

PR 25-MAY-2000; 2000US-207454P.

XX

PR 09-JUN-2000; 2000US-211314P.

XX

PR 18-JUL-2000; 2000US-219007P.

XX

PR 13-DEC-2000; 2000US-255281P.

XX

(MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

XX

Schlegel R, Endege WO, Monahan JE;

XX

WPI: 2001-662795/76.

XX

Novel isolated nucleic acid molecule associated with cancerous state of prostate cells and correlating with presence of prostate cancer, useful for detecting presence of prostate cancer, stage of prostate cancer -

XX

Claim 1; Page 3869; 11750pp; English.

XX

The invention relates to an isolated nucleic acid molecule (I) comprising a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the specification or its complement. (I) is useful for:

XX

(a) assessing whether a patient is afflicted with prostate cancer;

XX

(b) monitoring the progression of prostate cancer in a patient;

XX

(c) assessing the efficacy of a test compound to inhibit prostate cancer in a patient;

XX

(d) assessing the efficacy of a therapy for inhibiting prostate cancer in a patient;

XX

(e) selecting a composition for inhibiting prostate cancer in a patient;

XX

(f) assessing the prostate cell carcinogenic potential of a compound;

XX

(g) determining whether prostate cancer has metastasized in a patient;

XX

(h) assessing the aggressiveness or indolence of prostate cancer in a patient;

XX

(I) is also useful as a pharmacodynamic or pharmacogenomic marker.

XX

RESULT 43

ABV28163

ID ABV28163 standard; cDNA; 1201 BP.

XX

AC ABV28163;

XX

DT 16-SEP-2002 (first entry)

XX

DE Human prostate expression marker cDNA 28154.

XX

KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;

XX

pharmacogenomic marker; gene; ss.

XX

OS Homo sapiens.

XX

PN WO200160860-A2.

XX

PD 23-AUG-2001.

XX

PF 20-FEB-2001; 2001WO-US05171.

XX

PR 17-FEB-2000; 2000US-183319P.

XX

PR 16-MAR-2000; 2000US-189862P.

XX

PR 25-MAY-2000; 2000US-207454P.

XX

PR 09-JUN-2000; 2000US-211314P.

XX

PR 18-JUL-2000; 2000US-219007P.

XX

PR 13-DEC-2000; 2000US-255281P.

XX

(MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

XX

Schlegel R, Endege WO, Monahan JE;

XX

WPI: 2001-662795/76.

XX

Novel isolated nucleic acid molecule associated with cancerous state of prostate cells and correlating with presence of prostate cancer, useful for detecting presence of prostate cancer, stage of prostate cancer -

XX

Claim 1; Page 5839-5840; 11750pp; English.

XX

The invention relates to an isolated nucleic acid molecule (I) comprising a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the specification or its complement. (I) is useful for:

XX

(a) assessing whether a patient is afflicted with prostate cancer;

XX

(b) monitoring the progression of prostate cancer in a patient;

XX

(c) assessing the efficacy of a test compound to inhibit prostate cancer in a patient;

XX

(d) assessing the efficacy of a therapy for inhibiting prostate cancer in a patient;

XX

(e) selecting a composition for inhibiting prostate cancer in a patient;

XX

(f) assessing the prostate cell carcinogenic potential of a compound;

XX

(g) determining whether prostate cancer has metastasized in a patient;

XX

(h) assessing the aggressiveness or indolence of prostate cancer in a patient;

XX

(I) is also useful as a pharmacodynamic or pharmacogenomic marker.

XX

SQ Sequence 1201 BP; 355 A; 291 C; 289 G; 265 T; 1 other;

XX

Query Match 13.6%; Score 18; DB 23; Length 1201;

XX

Best Local Similarity 100.0%; Pred. No. 23;

XX

Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

XX

QY 40 GATAAGGCCAAGCTGAAG 57

XX

|||||

XX

401 GATAAGGCCAAGCTGAAG 418

XX

RESULT 44

XX

ABK84119

XX

ID ABK84119 standard; cDNA; 1262 BP.

XX

AC ABK84119;

XX

DT 14-AUG-2002 (first entry)

XX

Query Match 13.6%; Score 18; DB 24; Length 1262;

Best Local Similarity 100.0%; Pred. No. 23;

Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 40 GATAAGCCCAAGCTGAAG 57

Db 603 GATAAGCCCAAGCTGAAG 620

RESULT 45

AAS02039

ID AAS02039 standard; cDNA; 1517 BP.

XX AC AAS02039;

XX DT 16-JUL-2001 (first entry)

XX DE DNA encoding molecule for disease detection and treatment, mddt4.

XX KW Human; mddt4; zinc finger; gene therapy; adenosine deaminase deficiency;

XX KW ADA; severe combined immunodeficiency syndrome; cystic fibrosis;

XX KW thalassemia; familial hypercholesterolemia; haemophilia; factor VIII;

XX KW factor IX; cancer; cell proliferation; parasite; human retrovirus; HIV;

XX KW hepatitis B; hepatitis C; Candida albicans; Plasmodium falciparum;

XX KW Paracoccidioides brasiliensis; Trypanosoma brasiliensis; ss.

XX OS Homo sapiens.

XX PN WO200123538-A2.

XX PD 05-APR-2001.

XX PF 22-SEP-2000; 2000WO-US26085.

XX PR 28-SEP-1999; 99US-0158565.

XX PR 30-NOV-1999; 99US-0168197.

XX PA (INCY-) INCYTE GENOMICS INC.

XX PI Hodgson DM, Lincoln SE, Russo FD, Spiro PA, Banville SC;

XX PI Bratcher SR, Dufour GE, Cohen HJ, Rosen BH, Chalup MS;

XX PI Hillman JL, Jones AL, Yu JY, Greenwalt LB, Panzer SR;

XX PI Roseberry AM, Wright RJ, Chen W, Liu TF, Yap PE, Stockdreher TK;

XX PI Anshey S, Fong WT;

XX DR WPI; 2001-258131/26.

XX PT Purified disease treatment and detection molecule polynucleotides and

XX PT polypeptides, useful for providing diagnostic assays and gene therapy -

XX PS Claim 1; Page 97-98; 113pp; English.

XX CC The sequence represents the coding sequence of molecule for disease
XX CC detection and treatment, mddt4, shown by computer analysis to be similar
XX CC to zinc finger C2H2 type family of proteins. The sequence may be used for
XX CC somatic or germline gene therapy. Gene therapy may be performed to: (i)
XX CC correct genetic deficiency such as in severe immunodeficiency syndrome
XX CC associated with adenosine deaminase (ADA) deficiency, cystic fibrosis,
XX CC thalassemias, familial hypercholesterolemia and haemophilia caused by
XX CC factor VIII or factor IX deficiencies; (ii) express a conditional lethal
XX CC gene product (such as in the case of cancers which result from
XX CC unregulated cell proliferation); (iii) express a protein which affords
XX CC protection against intracellular parasites (for example, human
XX CC retroviruses such as HIV, hepatitis B or C, fungal parasites such as
XX CC Candida albicans and Paracoccidioides brasiliensis, and protozoal
XX CC parasites such as Plasmodium falciparum and Trypanosoma brasiliensis.
XX SQ Sequence 1517 BP; 289 A; 544 C; 447 G; 233 T; 4 other;

Query Match 13.6%; Score 18; DB 22; Length 1517;

Best Local Similarity 100.0%; Pred. No. 23;

Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Human cDNA differentially expressed in granulocytic cells #690.

Human; ss; granulocytic cell; DNA chip; bacterial infection;

viral infection; parasitic infection; protozoal infection;

fungal infection; sterile inflammatory disease; psoriasis;

rheumatoid arthritis; glomerulonephritis; asthma; thrombosis;

cardiac reperfusion injury; renal reperfusion injury; ARDS;

adult respiratory distress syndrome; inflammatory bowel disease;

Crohn's disease; ulcerative colitis; periodontal disease;

granulocyte activation; chronic inflammation; allergy.

XX OS Homo sapiens.

XX WO200228999-A2.

XX 11-APR-2002.

XX 03-OCT-2001; 2001WO-US30821.

XX 03-OCT-2000; 2000US-237189P.

XX (GENE-) GENE LOGIC INC.

XX Beazer-Barclay Y, Weissman SM, Yamaga S, Vockley J;

XX WPI; 2002-435328/46.

XX Detecting granulocyte activation by detecting differential expression

XX of genes associated with granulocyte activation, which serves as

XX diagnostic markers that is useful for monitoring disease states and

XX drug toxicity -

XX Claim 1; SEQ ID NO 690; 114pp; English.

XX The invention relates to detecting (M1) granulocyte (GC) activation

XX (GCA); by detecting the level of expression of gene(s) (Gs) identified by

XX DNA chip analysis as given in the specification, and comparing

XX the expression level to an expression level in an unactivated

XX GC, where differential expression of Gs is indicative of GCA.

XX Also included are modulation of at least one gene in Gs; (2) screening (M3)

XX that alters the expression of GCA or an inflammation (especially

XX for an agent capable of modulating GCA or an inflammation (especially

XX chronic) in a tissue, an allergic response in a subject, exposure of a

XX subject to a pathogen or sterile inflammatory disease using the

XX gene expression profile; (3) detecting (M4) an inflammation (especially

XX chronic) in a tissue, an allergic response in a subject, exposure of a

XX subject to a pathogen or sterile inflammatory disease, by detecting the

XX level of expression in a sample of the tissue of gene(s) from Gs, where

XX the level of expression of the gene is indicative of inflammation;

XX (4) treating (M5) an inflammation (especially chronic) or in a tissue,

XX an allergic response in a subject, exposure of a subject to a pathogen

XX or sterile inflammatory disease, by contacting a tissue having

XX inflammation with an agent that modulates the expression of gene(s)

XX from Gs in the tissue. M1 is useful for detecting GCA; M2 is useful for

XX modulating GCA; M3 is useful for screening an agent capable of modulating

XX GCA preferably in an inflammation (especially chronic) in a tissue, an allergic

XX response in a subject, exposure of a subject to a pathogen or sterile

XX inflammatory disease (e.g. psoriasis, rheumatoid arthritis, renal

XX glomerulonephritis, asthma, thrombosis, cardiac reperfusion injury, renal

XX reperfusion injury, ARDS, adult respiratory distress syndrome,

XX inflammatory bowel disease, Crohn's disease, ulcerative colitis,

XX parasitic infection, protozoal infection, fungal infection, and M5 is

XX useful for treating one of the above conditions. The present

XX sequence represents a gene differentially expressed in granulocytes.

XX Note: The sequence data for this patent did not form part

XX of the printed specification, but was obtained in electronic

XX format directly from WIPO at

XX ftp.wipo.int/pub/published_pat_sequences.

XX Sequence 1262 BP; 240 A; 404 C; 400 G; 215 T; 3 other;

QY 43 AAGCCCAAGCTGAAGGCC 60
XXXXXXXXXXXXXXXXXXXX
Db 50 AAGCCCAAGCTGAAGGCC 67
XXXXXXXXXXXXXXXXXXXX

RESULT 46
ABL06285
ID ABL06285 standard; cDNA; 2606 BP.
XX AC ABL06285;
XX XX
XX 26-MAR-2002 (first entry)
DT XX
XX Drosophila melanogaster expressed polynucleotide SEQ ID NO 13337.
DE XX
XX Drosophila; developmental biology; cell signalling; insecticide;
KW KW
KW pharmaceutical; gene; ss.
XX XX
OS Drosophila melanogaster.
XX WO200171042-A2.
XX XX
XX 27-SEP-2001.
XX XX
XX 23-MAR-2001; 2001WO-US09231.
XX XX
PR 23-MAR-2000; 2000US-191637P.
PR 11-JUL-2000; 2000US-0614150.
XX XX
PA (PEKE) PE CORP NY.
XX XX
PI Venter JC, Adams M, Li PWD, Myers EW;
XX WPI; 2001-656860/75.
XX P-PSDB; ABB62182.
XX XX
PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -
XX XX
XX Claim 1; SEQ ID NO 13337; 21pp + Sequence Listing; English.
XX XX
CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins
CC (ABB57737-ABB72072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX XX
XX Sequence 2606 BP; 711 A; 667 C; 633 G; 595 T; 0 other;
XX XX
Query Match 13.6%; Score 18; DB 23; Length 2606;
Best Local Similarity 100.0%; Pred. No. 23;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 44 AGGCCCAAGCTGAAGGCCA 61
XXXXXXXXXXXXXXXXXXXX
Db 1712 AGGCCCAAGCTGAAGGCCA 1729
XXXXXXXXXXXXXXXXXXXX

RESULT 47
ABL21125
ID ABL21125 standard; DNA; 2609 BP.
XX AC ABL21125;
XX XX
XX 26-MAR-2002 (first entry)
DT XX
XX XX

DE Drosophila melanogaster genomic polynucleotide SEQ ID NO 14848.
XX XX
XX Drosophila; developmental biology; cell signalling; insecticide;
KW KW
KW pharmaceutical; gene; ds.
XX XX
OS Drosophila melanogaster.
XX WO200171042-A2.
XX XX
XX 27-SEP-2001.
XX XX
XX 23-MAR-2001; 2001WO-US09231.
XX XX
PR 23-MAR-2000; 2000US-191637P.
PR 11-JUL-2000; 2000US-0614150.
XX XX
PA (PEKE) PE CORP NY.
XX XX
PI Venter JC, Adams M, Li PWD, Myers EW;
XX WPI; 2001-656860/75.
XX XX
PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -
XX XX
XX Claim 1; SEQ ID NO 14848; 21pp + Sequence Listing; English.
XX XX
CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins
CC (ABB57737-ABB72072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX XX
XX Sequence 2609 BP; 681 A; 670 C; 656 G; 602 T; 0 other;
XX XX
Query Match 13.6%; Score 18; DB 23; Length 2609;
Best Local Similarity 100.0%; Pred. No. 23;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 44 AGGCCCAAGCTGAAGGCCA 61
XXXXXXXXXXXXXXXXXXXX
Db 1715 AGGCCCAAGCTGAAGGCCA 1732
XXXXXXXXXXXXXXXXXXXX

RESULT 48
AAK52974
ID AAK52974 standard; cDNA; 3291 BP.
XX AC AAK52974;
XX XX
XX 06-NOV-2001 (first entry)
DT XX
XX Human polynucleotide SEQ ID NO 2503.
DE XX
XX Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW tissue growth factor; immunomodulatory; cancer; leukaemia;
KW nervous system disorder; arthritis; inflammation; ss.
XX XX
OS Homo sapiens.
XX XX
XX WO200157190-A2.
XX PN
XX 09-AUG-2001.
XX PD
XX 05-FEB-2001; 2001WO-US04098.

us-09-915-178-1.olil0.rng

Wed Jun 4 11:04:53 2003

```
XX 03-FEB-2000; 2000US-0496914.
PR 27-APR-2000; 2000US-0560875.
PR 20-JUN-2000; 2000US-0598075.
PR 19-JUL-2000; 2000US-0620325.
PR 01-SEP-2000; 2000US-0654936.
PR 15-SEP-2000; 2000US-0663561.
PR 20-OCT-2000; 2000US-0693325.
PR 30-NOV-2000; 2000US-0728422.
XX (HYSE-) HYSEQ INC.
XX Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y, Ma Y;
PI Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;
PI Xue AJ, Yang Y, Wejhrman T, Goodrich R;
XX WPI: 2001-476283/51.
DR P-PSDB; AAM79841.
XX Nucleic acids encoding polypeptides with cytokine-like activities,
DR useful in diagnosis and gene therapy -
XX Claim 1; Page 4781-4782; 6221pp; English.
XX The invention relates to polynucleotides (AAK51456-AAK53435) and the
XX encoded polypeptides (AAM78323-AAM80302) that exhibit activity elating to
XX cytokine, cell proliferation or cell differentiation or which may induce
XX production of other cytokines in other cell populations. The
XX polynucleotides and polypeptides are useful in gene therapy, vaccines or
XX peptide therapy. The polypeptides have various cytokine-like activities,
XX e.g. stem cell growth factor activity, haematopoiesis regulating
XX activity, tissue growth factor activity, immunomodulatory activity and
XX activin/inhibin activity and may be useful in the diagnosis and/or
XX treatment of cancer, leukaemia, nervous system disorders, arthritis and
XX inflammation.
XX Note: Records for SEQ ID NO 2110 (AAK52581), 2111 (AAK52582) and 3666
XX (AAM80020) are omitted as the relevant pages from the sequence listing
XX were missing at the time of publication.
XX Sequence 3291 BP; 620 A; 1131 C; 989 G; 551 T; 0 other;
XX Query Match 13.6%; Score 18; DB 22; Length 3291;
XX Best Local Similarity 100.0%; Pred. No. 23;
XX Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 44 AGGCCAAGCTGAAGGCCA 61
DB 2929 AGGCCAAGCTGAAGGCCA 2946
RESULT 49
AAK51990
ID AAK51990 standard; cDNA; 3754 BP.
XX AC AAK51990;
XX DE 06-NOV-2001 (first entry)
XX DE Human polynucleotide SEQ ID NO 535.
XX Human; cytokine; cell proliferation; cell differentiation; gene therapy;
XX vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
XX tissue growth factor; immunomodulatory; cancer; leukaemia;
XX nervous system disorder; arthritis; inflammation; ss.
XX Homo sapiens.
XX WO200157190-A2.
XX 09-AUG-2001.
XX 05-FEB-2001; 2001WO-US04098.
XX
XX 03-FEB-2000; 2000US-0496914.
PR 27-APR-2000; 2000US-0560875.
PR 20-JUN-2000; 2000US-0598075.
PR 19-JUL-2000; 2000US-0620325.
PR 01-SEP-2000; 2000US-0654936.
PR 15-SEP-2000; 2000US-0663561.
PR 20-OCT-2000; 2000US-0693325.
PR 30-NOV-2000; 2000US-0728422.
XX (HYSE-) HYSEQ INC.
XX Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y, Ma Y;
PI Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;
PI Xue AJ, Yang Y, Wejhrman T, Goodrich R;
XX WPI: 2001-476283/51.
DR P-PSDB; AAM78857.
XX Nucleic acids encoding polypeptides with cytokine-like activities,
DR useful in diagnosis and gene therapy -
XX Claim 1; Page 1931-1934; 6221pp; English.
XX The invention relates to polynucleotides (AAK51456-AAK53435) and the
XX encoded polypeptides (AAM78323-AAM80302) that exhibit activity elating to
XX cytokine, cell proliferation or cell differentiation or which may induce
XX production of other cytokines in other cell populations. The
XX polynucleotides and polypeptides are useful in gene therapy, vaccines or
XX peptide therapy. The polypeptides have various cytokine-like activities,
XX e.g. stem cell growth factor activity, haematopoiesis regulating
XX activity, tissue growth factor activity, immunomodulatory activity and
XX activin/inhibin activity and may be useful in the diagnosis and/or
XX treatment of cancer, leukaemia, nervous system disorders, arthritis and
XX inflammation.
XX Note: Records for SEQ ID NO 2110 (AAK52581), 2111 (AAK52582) and 3666
XX (AAM80020) are omitted as the relevant pages from the sequence listing
XX were missing at the time of publication.
XX Sequence 3754 BP; 811 A; 1165 C; 1035 G; 743 T; 0 other;
XX Query Match 13.6%; Score 18; DB 22; Length 3754;
XX Best Local Similarity 100.0%; Pred. No. 23;
XX Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 44 AGGCCAAGCTGAAGGCCA 61
DB 3392 AGGCCAAGCTGAAGGCCA 3409
RESULT 50
ABL25848
ID ABL25848 standard; DNA; 4229 BP.
XX AC ABL25848;
XX DE 26-MAR-2002 (first entry)
XX DE Drosophila melanogaster genomic polynucleotide SEQ ID NO 29017.
XX Drosophila; developmental biology; cell signalling; insecticide;
XX pharmaceutical; gene; ds.
XX Drosophila melanogaster.
XX WO200171042-A2.
XX 27-SEP-2001.
XX 23-MAR-2001; 2001WO-US09231.
XX 23-MAR-2000; 2000US-191637P.
XX 11-JUL-2000; 2000US-0614150.
XX
```


PA (PEKE) PE CORP NY.
XX
PI Venter JC, Adams M, Li PWD, Myers EW;
XX
DR WPI; 2001-656860/75.
XX
XX New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -
XX
PS Claim 1; SEQ ID NO 29017; 21pp + Sequence Listing; English.
XX
XX The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins
CC (ABB57737-ABB72072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 4229 BP; 1287 A; 1022 C; 928 G; 992 T; 0 other;

Query Match 13.6%; Score 18; DB 23; Length 4229;
Best Local Similarity 100.0%; Pred. No. 23;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 29 TTGCCAGCTTGGATAAGG 46
|||||
Db 2080 TTGCCAGCTTGGATAAGG 2097

Search completed: June 2, 2003, 15:56:08
Job time : 250 secs

THIS PAGE BLANK (3370)